

Effects of a summer camp program on enhancing the academic achievement
skills of children with Autism Spectrum Disorders.

A dissertation submitted to the
Graduate school in partial fulfillment of the
requirements for the degree
of doctor of philosophy

By

Teresa L. Davenport

Dissertation advisor: Dr. David McIntosh

Ball State University
Muncie, IN

Acknowledgement

I owe the deepest gratitude for this dissertation to my mentor professor, Dr. David McIntosh. Throughout my academic career, Dr. McIntosh provided numerous opportunities for me to develop my professional skills both within an academic setting as well as in a clinical setting. Through his guidance and instruction, Dr. McIntosh has shaped my academic and clinical careers, for which I am forever grateful.

In addition to being forever grateful to Dr. Theresa Kruczek for her guidance while writing this dissertation, I will always remember her mentorship of me during my counseling classes and practicum. Her talent for providing therapy to children and adolescents is of a caliber I will strive to reach throughout my career. As a result of her talented instruction, I hope to pass along her wisdom and emulate her kindness with my future clients.

I am deeply indebted to Dr. Holmes Finch, for without a doubt this dissertation would not have been possible without his assistance. However, what I am most grateful to Dr. Finch for was his always gentle and patient instruction throughout my grad school career; it was of great benefit to me during the repeated explanations of statistical operations I so often required. I am confident I will be a better teacher of my future students as a result of having him as a role model.

I would also like thank Dr. Catherine Chen for being so flexible with meeting with me and thorough editing of the drafts of my dissertation. In addition, a special thanks to Penny Nowaczyk and Linda Jones for all their assistance the past few years with scheduling meetings and classes, filing paperwork, letting me shamelessly abuse my copy machine privileges, and all that delicious cake.

All my fellow graduate students deserve my gratitude, as they have all come to feel like a second family for me, which was a constant comfort to me throughout my graduate career. However, I owe special thanks to a few fellow students in particular for their support of me. To Renee Grizzle, for all those long nights spent working on our dissertations together at the MT Cup, and for being an excellent travel companion and emotional support through the process of applying for internship. To Amy Zimmerman, for reminding me that that everything is easier if you smile while doing it. To Anya Mazur-Mosiewicz for inspiring me to always work a little bit harder, and to Jenny Vought, for reminding me that no matter how hard I work, people will also remember me for what I bring to the party. And to Matt Holcomb, for soothing words that sought to keep me grounded during my frequent bouts of anxiety, and for providing me the occasional liquid assistance with relaxing when that was also necessary.

Dedication

To my mother, Nancy Davenport,
for her constant encouragement and support throughout graduate school.

Abstract

The purpose of this study was to determine if a summer camp utilizing academic and behavioral remediation programming could increase the academic achievement of children with autism spectrum disorders. Academic achievement was measured using the Wide Range Achievement Test-Fourth Edition (WRAT4; Wilkinson & Robertson, 2006) and an Informal Reading Inventory (IRI; Burns & Roe, 2002) in a sample of 21 children ranging from 6 years, 0 months to 13 years, 1 month prior to and after attending a summer camp for children with autism around a Midwestern city. Repeated measures MANOVAs were used to analyze WRAT4 and IRI assessment results. MANOVA results indicated there were no significant changes on any of the tests of the WRAT4 over time after a Bonferroni correction. MANOVA results further indicated there was no interaction between gender and time. A nonparametric sign test also was conducted to examine changes in the WRAT4 tests. Results of the nonparametric sign test did not reveal significant changes in the WRAT4 tests; however, the Sentence Comprehension test did approach significance. Although participants did not make statistical gains on the WRAT4, they also did not demonstrate a decline in academic skills as measured by the WRAT4. Participants demonstrated less consistent results on the word lists of the IRI as well as the passage comprehension component of the IRI. The number of children who declined or improved on the IRI overtime appeared to be random. A qualitative analysis of four participants selected based on age and gender did not reveal any hypotheses for the differences in performance on the IRI over time. Suggestions for future research include replicating the study with a larger sample, children with multicultural backgrounds, and controlling for the cognitive abilities of the participants.

TABLE OF CONTENTS

ACKNOWLEDGEMENT	ii
DEDICATION	iii
ABSTRACT	iv
LIST OF TABLES	viii
LIST OF FIGURES	ix
CHAPTER I: INTRODUCTION.....	1
Educational needs of children with autism	1
Overview of autism	2
Adventure Therapy	3
Therapeutic summer camps for children with disabilities	4
Therapeutic summer camps for children with autism	7
Academic summer programs	10
Wide Range Achievement Test, Fourth Edition (WRAT4)	11
Informal Reading Inventories (IRIs).....	13
Academic skills in children with autism	14
Purpose of the study.....	18
Significance of the study.....	19
CHAPTER II: REVIEW OF THE LITERATURE	22
Summer camps for children with disabilities.....	22
Academic summer programs	31
Academic achievement of children with autism	36
Reading	36
Written Language.....	40
Mathematics	41
Wide Range Achievement Test	43
Description.....	43
History.....	43
Psychometric Properties.....	44
Validity	45
Reliability.....	47
Informal Reading Inventory.....	49
Description.....	49

History.....	51
Psychometric Properties.....	54
Validity	54
Reliability.....	59
Summer camps for children with autism	62
Summary	69
 CHAPTER III: RESEARCH METHODOLOGY	 71
Participants.....	71
Procedures	72
Setting	72
Camp Format	74
Academic Programming	76
Examiner Training	77
Assessment Administration	77
Instrumentation	78
Wide Range Achievement Test – Fourth Edition (WRAT4).....	79
Description.....	79
Standardization	80
Validity	80
Reliability.....	82
Informal Reading Inventory (Burns & Roe, 2002).....	84
Description.....	84
Test Construction.....	84
Descriptive Statistics.....	86
Research Design.....	86
Data analysis	87
Sample Size.....	88
Fit Statistics.....	88
 CHAPTER IV: RESULTS.....	 89
Research Hypothesis	89
Descriptive Statistics.....	89
MANOVA results	91
Sign Test	91
Research Question	92
Descriptive Statistics.....	92
Qualitative Analysis.....	94
Analysis of Error Rates	99

CHAPTER V: DISCUSSION	103
Rational for the study	103
Research Hypothesis	104
Research Question	108
Limitations	116
Suggestions for future research	119
References	123
APPENDIX A: Parent Consent Forms and Demographic Questionnaire	142
APPENDIX B: IRB approval	148

List of Tables

Table 3.2	<i>Description of WRAT4 tests</i>	80
Table 4.1	<i>Descriptive Statistics for Standard Scores on the Wide Range Achievement Test – Fourth Edition (WRAT4)</i>	90
Table 4.2	<i>Sign Test Results for first and second administrations of tests on the Wide Range Achievement Test – Fourth Edition (WRAT4)</i>	92
Table 4.3	<i>Number of students who improved or declined as measured by the word list for the IRI</i>	93
Table 4.4	<i>Number of students who improved or declined as measured by Passage Comprehension for the IRI</i>	94
Table 4.5	<i>Means and standard deviations of error rates for each question type on the IRI</i>	100
Table 4.6	<i>Paired samples t-test for error rates on IRI between first and second administration</i>	101

List of Figures

Figure 4.1	WRAT4 and IRI scores for Child 1	95
Figure 4.2	WRAT4 and IRI scores for Child 2	96
Figure 4.3	WRAT4 and IRI scores for Child 3	97
Figure 4.4	WRAT4 and IRI scores for Child 4	98

CHAPTER I

Introduction

This section briefly reviews the literature relevant to the current study and provides rational for the study. This section also outlines the purpose of the study and indicates the significance of this study to the field of psychology.

Educational needs of children with autism

Children with autism may be more likely than their typically developing peers to lose skills they do not practice regularly (Heflin & Alaimo, 2007; Van Wert & Reitz, 1978). Although children with autism may be able to regularly practice academic skills during the school year, public school systems typically do not provide services for students through summer months. The discontinuation of services through the summer may interrupt the development of academic skills in children with autism.

The National Research Council (2001) concluded that effective, comprehensive intervention programs for children with autism utilize intensive programming throughout the year. The Individuals with Disabilities Education Improvement act (IDEA; 2004) states that a free and appropriate education (FAPE) must be provided to all children with disabilities during the academic school year, but does not mandate the provision of services to all students with disabilities during the summer. IDEA also states that school corporations may be required to provide services to students with disabilities during the summer months only if a case conference committee determines it is necessary in order to provide a FAPE to the student. However, IDEA does not clearly stipulate methods a case conference must use to determine if summer instruction would be necessary. Legal precedent used in the past has involved a documentation that regression will occur over long periods of time away from school and that recoupment of skills

will require more than a reasonable amount of time (Boomer & Garrison-Harrell, 1995; Mandlawitz, 2005). That is, although a child may continue to make significant gains if provided special education services through the summer months, the school may not be legally obligated to provide services. Because parents may not be able to obtain educational services during the summer through their local public school, many have sought other opportunities for their children to receive services during the summer months (Brookman et al. 2003). Many parents choose summer camps as a way to maintain their children's academic skills.

Overview of Autism

Autism is characterized as deficits in social interaction and communication, as well as the presence of stereotyped and restricted patterns of behavior and interests. Autism is classified as a Pervasive Developmental Disorder (PDD), a category that includes Asperger's Disorder (AD) and Pervasive Developmental Disorder – Not Otherwise Specified (PDD-NOS) (American Psychiatric Association [APA], 2000). While psychologists tend to make distinctions between the different subtypes of autism, it is common for school officials and other agencies to use the term Autism Spectrum Disorders (ASDs) to refer to autism, Asperger's Disorder, and PDD-NOS, due to similarities between the disorders. Criteria for Asperger's Disorder, sometimes referred to as Asperger's Syndrome, include the same deficits in social interaction and the presence of stereotyped and repetitive interests as autism, and may demonstrate deficits in the use of language such as struggling with pragmatic language and the presence of pedantic speech, but there is no history of the overall language delay as seen in autism. A diagnosis of PDD-NOS is made when a child demonstrates deficits in the areas specific to autism, but the deficits in all or some areas are subthreshold and therefore, a diagnosis of a specific PDD cannot be made (APA, 2000).

To be diagnosed with autism an individual must show characteristics of the disorder within the first three years of life (APA, 2000). Characteristics of other PDDs may be recognized later than three years of age, but all PDDs have an onset during childhood or adolescence (APA, 2000). There is now evidence that an accurate clinical diagnosis of autism can be made as early as 2-years-of-age (Coonrod & Stone, 2005; Lord, 1995).

Studies examining the trends of incidence and prevalence rates of autism from 1966 to 2003 have not provided strong evidence to support the widely held belief among the general public that there has been an increase in the incidence of autism (Fombonne, 2005). However, there is evidence that the prevalence rate of autism is increasing. The specific reasons for the increase in prevalence are not known, though likely causes include improvements in diagnostic techniques, changes in the policies for special education, and the increasing availability of services (Fombonne, 2005). At least one computer simulation has been used to demonstrate that a broadening of diagnostic criteria for Autistic Disorder could account for increases in prevalence rates (Wazana, Bresnahan & Kline, 2007). Current estimates for the prevalence of autism are around 0.6% for the general population (Fombonne, 2005; Rice, 2007).

Due to the increasing prevalence rate of autism, and because diagnoses of autism are typically made in childhood and adolescence, public school systems have seen an increase in enrollment of children with autism. Thus, school systems have had an increased focus on how best to provide educational services to children with autism, including a careful examination of their legal obligations for provision of services.

Adventure Therapy

Adventure therapy and the related forms of therapy (e.g., wilderness therapy, therapeutic camping, low-ropes courses, experiential education) are generally described as forms of therapy

using nontraditional settings and activities as vehicles for providing therapeutic interventions to individuals and groups (Davis-Berman & Berman, 1994; Glass & Myers, 2001; Russell, 2001). Adventure therapy techniques are particularly well suited for use with individuals who may not respond well to traditional talk therapies, including typically developing children and adolescents (Davis-Berman & Berman, 1994). Similarly, due to the verbal and communication deficits that are hallmark symptoms of autism spectrum disorders, it is logical to consider the delivery of therapeutic services to children with autism spectrum disorders in nontraditional settings, such as summer camps, which are used in adventure therapy. Summer camps, typically constructed around structured, hands-on activities and projects rather than verbal discussions, may provide an excellent platform to deliver therapeutic services to children with autism. In addition, the use of settings such as summer camps necessarily involve group activities and interactions with other campers and staff, providing an opportunity for social skills development and intervention that may not be available in traditional, individual therapy settings.

Therapeutic summer camps for children with disabilities

Summer camps have been used to deliver therapeutic mental health services to individuals of all ages with several types of disabilities, including physical disabilities (Briery & Rabian, 1999; Eng & Davies, 1991; Jelalian, Mehlenbeck, Lloyd-Richardson, Birmaher & Wing, 2006; Klee, Greenleaf & Watkins, 1997; Martiniuk, 2003; Stevens, 2004), as well as psychological disorders (e.g., Learning Disabilities, Attention-Deficit Hyperactivity Disorder, Developmental Disabilities) (Bennett, Cardone & Jarczyk, 1998; Davis-Berman & Berman, 1989; Eikeanaes, Gude & Hoffart, 2006; Forehand, Mulhern & Rickard, 1969; Gillis & Simpson, 1991; Heckel, Hursh & Hiers, 1977; Henley, 1999, Hinshaw, 2002; Hung & Thelander, 1978; Kelley, Coursey & Selby, 1997; Larson, 2007; Lattal, 1969; Romi & Kohan, 2004; Russell,

2003; Stich & Senior, 1984; Yssel, Margison, Cross & Merbler, 2005). The structure of the camps, including their curriculum and the composition of the camper population, depends greatly on the nature of the camp and the therapeutic goals of those who design the camps.

Some summer camps have utilized a specific, therapeutic curriculum to treat a specific disorder. For example, a summer camp for children with diabetes may include educational sessions regarding medication monitoring, using equipment, and healthy lifestyles – activities that directly impact diabetes – in addition to traditional camp activities such as swimming and arts and crafts (Briery & Rabian, 1999). More commonly, however, adventure therapies and summer camps have utilized traditional outdoor or camp experiences and have extrapolated therapeutic lessons from such activities (Gillis & Simpson, 1991; Jones, 2004; Kelley et al. 1997; Larson, 2007; Marx, 1988). For example, one program treating people who abuse substances utilized a ropes course to teach skills such as teamwork and reliance on others for successful completion of the course, linking these skills to the use of interventions such as using a 12-step program in order to remain successful in maintaining sobriety. This program resulted in 27% higher rates of sobriety 10 months after treatment compared to individuals who did not complete the ropes course (Bennett et al. 1998). Some programs combine direct and indirect treatments. For example, one such camp for people who abuse substances utilized formal interventions such as lectures on drug abuse combined by camping experiences designed to foster cooperation and trust. This program found statistically significant reductions in both objective and self-ratings of problem behaviors, as well as a reduction on several clinical scales on the Minnesota Multiphasic Personality Inventory (MMPI) (Gillis & Simpson, 1991).

One method of providing a summer camp experience to children with disabilities is to include children with disabilities in camps designed for, and consisting of, primarily typically

developing peers (Blas, 2007; Brookman et al. 2003). Some of these camps included children with many types of disabilities, and so they used a variety of interventions specific to certain disorders for each camper to successfully incorporate each camper into the camp (Blas, 2007). In another example, one summer camp specifically targeted campers with autism and correspondingly only used interventions specific to autism. Results of this study did not utilize formal measures, but reported campers with autism were able, with the assistance of aides, to complete all camp activities and indicated typically developing campers as well as staff members benefited from the inclusion of campers with autism (Brookman et al. 2003).

Alternatively, segregated summer camps designed exclusively for children with disabilities may offer therapeutic benefits to campers that cannot be found in inclusive camps. Individuals attending camps designed exclusively for individuals with disabilities have been found to produce a feeling of belonging and a chance to develop meaningful relationships with peers who share similar difficulties (Michalski, Mishna, Worthington & Cummings, 2003; Yssel et al. 2005). In addition, segregated summer camps encourage independence and a chance for an individual to challenge the limitations of his or her disability without the feeling of constant comparison to typically developing peers (Goodwin & Staples, 2005). Although children with disabilities, including autism, may benefit from having typically developing peers for role models, a segregated summer camp offers a child with autism the opportunity to interact on a meaningful level with other campers who have, and therefore have an understanding of, the same social, emotional, and behavioral difficulties he or she has had. In addition, a segregated camp offers the opportunity to implement therapies with all campers, rather than singling out only campers with disabilities, which reduces the stigma of those therapies.

Many summer camps for children with psychiatric disorders have demonstrated effectiveness in reducing clinical symptoms and maladaptive behaviors (Briery & Rabian, 1999; Eng & Davies, 1991; Forehand et al. 1969; Henley, 1999; Hinshaw, 2002; Larson, 2007; Lattal, 1969; Michalski et al. 2003; Yssel et al. 2005). For example, parent reports regarding a one week long summer camp designed for children who were gifted and learning disabled (“twice exceptional”) tailored to their unique needs indicated the parents felt the camp produced positive increases in the self-esteem of the campers (Yssel et al. 2005). In another example, the frequency of disruptive behaviors of two children was significantly reduced through the use of token reinforcement delivered at summer camp for children with emotional disabilities (Forehand et al. 1969). At another camp, and objective measure, the Piers-Harris Children’s Self-Concept Scale (PHCSCS), was used to determine that a summer camp program aimed at reducing behavioral problems in adolescents was more effective ($p = 2.66$) at increasing the self-concept of the campers compared to adolescents who recieved more traditional therapies from a social service agency ($p = 1.06$) (Larson, 2007).

Therapeutic summer camps and summer programs for children with autism

Currently, little empirical research on summer camps designed exclusively for children with autism exists. Three summer programs targeting only children with autism (Hung & Thelander, 1978; Lopata, Thomeer, Volker & Nida, 2006; Lopata, Thomeer, Volker, Nida & Lee, 2008; Van Wert & Reitz, 1978) were found. Of these, only one of the summer programs (Hung and Thelander, 1978) occurred in a traditional summer camp setting (e.g., in the wilderness). Another summer program was held on a university campus and targeted children with Asperger’s Disorder and high-functioning autism (Lopata, Thomeer, Volker & Nida, 2006; Lopata, Thomeer, Volker, Nida & Lee, 2008). A third program run on a secondary school

campus focused on children with autism only (Van Wert & Reitz, 1978). These programs primarily focused on social skills (Lopata et al. 2006; Lopata et al. 2008) and behavior (Hung & Thelander, 1978; Van Wert & Reitz, 1978), rather than academic skills such as reading, writing, and mathematics.

During the first two years of their summer program, Lopata et al. (2006) included 21 children with Asperger's Disorder who attended the program six hours per day during the week, for a total of six weeks over the summer months. The next two years of the program, Lopata et al. (2008) expanded the program to include 54 children with Asperger's Disorder, high-functioning autism, and PDD-NOS. Curriculum for all four years of the program was based on *Skillstreaming*, a program which utilizes systematic procedures for teaching, modeling, role-playing, providing performance feedback, and the transfer of learning in the area of social skills. The program primarily followed a schedule in which the children completed four 70-minute treatment cycles during each day, with time allotted for transitions and lunch. Cycles included 20 minutes of structured social skills instruction followed by a 50-minute activity which required social interaction. The program utilized objective ratings completed by parents and staff members of the program to evaluate gains in social skills. Pre-test ratings were collected immediately from parents and after eight days of the program from staff members, ratings were then also collected at the program.

Results from the first two years of the summer program indicated significantly greater social skills, adaptability skills, and fewer atypical behaviors, as measured by the parent Behavior Assessment System for Children, Second Edition (BASC-II). The same scales completed by staff members revealed significant improvements in social skills, but did not indicate improvement in adaptability or a reduction in atypical behaviors. The discrepancy

between these ratings may be due to many factors, especially the lack of familiarity the staff had when completing the BASC-II at the outset of the camp (Lopata et al. 2006).

The following two years of the program found similar results. All the participants were again rated using the BASC-II. Results indicated an increase in social skills, and parent ratings of atypicality did not show a change where staff ratings showed a slight increase. This study also looked at the withdrawing behaviors and found a significant decrease in withdrawing behaviors on parent and staff ratings. The Adaptive Skills Composite scales for both parents and staff indicated improvements in adaptive functioning. In addition, this study found no difference from highly structured feedback about the children's use of social skills compared to a more non-structured feedback method (Lopata et al. 2008).

Van Wert and Reitz (1978) report on a summer program for children with autism primarily focusing on increasing adaptive functioning and reducing maladaptive behaviors. The activities of the program focused on the aforementioned areas (e.g. communications group, fine motor training). The program ran for three years and although they did not report an exact number of participants, it was noted the program accepted between 15 and 30 children ages 3- to 18-years-of-age who had been diagnosed with autism. Children were assessed before and after the camp using the Behavior Characteristics Progression (1973) checklist, and a percentage of change in the behaviors were calculated for each child at the end of camp. This checklist assessed skill development in 24 behavioral areas with 50 behavioral objectives in each area (Van Wert & Reitz, 1978). The authors did not present outcomes for the children, so evaluation of the program's effectiveness was not possible.

Hung and Thelander's (1978) summer camp was designed specifically for children with autism. Eighteen children participated in a three-week residential summer camp located in a

traditional summer camp setting (e.g., wilderness). This camp utilized a one-to-one camper to counselor ratio and provided individualized behavioral therapy to all campers. Specifically, the camp utilized Applied Behavioral Analysis (ABA) techniques (e.g., food and token reinforcement, shaping) to reduce problem behaviors and to increase adaptive skills. The camp schedule included four daily classes lasting 45 minutes each, two of which contained an academic component: Language and special education, and 2 classes which did not: swimming, and social interaction. However, only behavioral goals were assessed and there was no measure of the impact of academic instruction on the campers' achievement skills. The camp demonstrated positive effects on the behavior of participants, with 79% of the children making gains of 15% or more in at least one domain (Hung & Thelander, 1978).

Most of the studies examining the use of summer camps to deliver interventions to children with autism have focused on behavioral treatments (Lopata et al. 2006; Lopata et al. 2008; Hung & Thelander, 1978). However, it is important to note that these behaviors often interfere with learning, and so addressing these behaviors and the effectiveness of reducing these behaviors is important to establish in conjunction with determining the ability of a summer camp to improve the academic skills of children with autism.

Academic summer programs

A meta-analytic review of 39 studies examining the effects of summer vacation on academic achievement indicated students may lose as much as one month of grade level-equivalent skills over the course of the typical summer break (Cooper, Nye, Charlton, Lindsay & Greathouse, 1996). In addition, a review of the Atlanta public school summer school program indicated students from low socioeconomic status (SES) lost more academic achievement skills during the summer, but this loss was significantly less for students who attended summer school

(Heyns, 1978). These reviews indicate the loss of academic skills over the summer occurs for many students, but may particularly harm those students traditionally considered at-risk due to a low SES status. Thus, a logical approach to combating summer losses has been in the creation of summer schools and summer programming with an academic focus in order to reduce the loss of academic skills. In addition, summer programs targeting at-risk individuals should be accessible to individuals with low SES.

In their meta-analytic review of summer school programs, Cooper, Charlton, Valentine and Muhlenbruck (2000) determined summer school programs are effective in increasing the academic skills of students in general. Cooper et al. (2000) also looked specifically at students with learning deficiencies, defined as “at-risk”, “underachievers or below grade level”, “failing or retained”, and “learning disabled or otherwise impaired”, and noted that these students in general made academic gains or maintained academic skills when enrolled in programs offering academic remediation over the summer. It was noted the effect sizes of the studies included in their review would generally be considered “small” to “average”, depending on the guidelines of interpretation being used (Cooper et al. 2000). Nonetheless, a small gain in academic skills or the maintenance of skills throughout the summer months should be considered a positive outcome when the alternative is a loss of academic skills. Thus, there exists clear evidence that summer school programming helps students maintain or make small gains in academic skills through academic instruction during the summer months.

Wide Range Achievement Test, Fourth Edition (WRAT4)

Currently, the most popular measures of achievement are comprehensive batteries, such as the Wechsler Individual Achievement Test – Second Edition (WIAT-II; Wechsler, 2001), and the Woodcock-Johnson Tests of Achievement, Third edition (WJ-ACH-III; Woodcock, McGrew

& Mather, 2001) are comprised of numerous subtests. These tests allow for the collection of extensive information regarding the achievement skills of individuals, all of which is useful in the differential diagnosis of specific learning disabilities. However, these batteries are time consuming to administer, and the breadth of information is unnecessary for a less intensive investigation of an individual's achievement, such as when establishing a baseline of achievement. A more concise measure of achievement can be found in the Wide Range Achievement Test, Fourth Edition (WRAT4; Wilkinson & Robertson, 2006).

Since the publication of the first edition of the Wide Range Achievement Test (WRAT; Jastak, 1945) through the current edition, all versions of the WRAT have been a tool used to assess the academic achievement of children and adults (Wilkinson & Robertson, 2006). The first version of the WRAT was developed to compliment results obtained on the Wechsler-Bellevue Scales, developed by David Wechsler. The next edition of the WRAT was the Wide Range Achievement Test-Revised (WRAT-R), published in 1978, and the Wide Range Achievement Test-Third Edition (WRAT3) followed this edition in 1993. The most current edition, the WRAT4, was published in 2006. The WRAT4 contains four subtests: Word Reading, Sentence Comprehension, Spelling, and Math Computation. The WRAT4 contains two parallel forms, Blue and Green, and a Combined Form consisting of all the items on both the Blue and Green Form (Wilkinson & Roberson, 2006).

Due to the recent publication The WRAT4, empirical research regarding the validity and reliability of the test was not available. However, the authors of the WRAT4 included much of the material found on the WRAT3, which has demonstrated good validity and reliability in empirical studies (Ashendorf, Jefferson, Green & Stern, 2009; Klimczak, Bradford, Burright & Donovanick, 2000; Smith, Smith & Smithson, 1995; Snelbacker, Wilkinson, Robertson & Glutting,

2001; Wilkinson & Robertson, 2006). Within the manual for the WRAT4, the authors report moderate to high correlations with cognitive tests and state these correlations are comparable to correlations found in those of other studies correlating cognitive abilities and achievement scores (Dell, Harrold & Dell, 2008; Wilkinson & Robertson, 2006). One study looking at the test-retest reliability of the reading test on the WRAT3 indicated a correlation of 0.90 over the course of one year within a geriatric population (Ashendorf et al. 2009). High levels of reliability have been reported for the WRAT4, including overall internal consistency ranging from 0.94-0.98, and alternate forms reliability of 0.88 for all ages (Dell et al. 2008; Wilkinson & Robertson, 2006). Thus, although empirical research on the psychometric properties of the WRAT4 is not currently available, available evidence suggests it has adequate validity and reliability.

Informal Reading Inventories (IRIs)

Informal Reading Inventories (IRIs) are reading assessments designed to measure an individual's grade level in reading. Typically, IRIs include reading levels which begin at the pre-primary level and go through 12th grade (Burns & Roe, 2002; Spector, 2005). IRIs offer an estimate of an individual's independent reading level, where he or she can read and understand with ease, instructional level, where they can read with understanding if he or she has assistance from a teacher, and frustration level, where the material is too difficult for him or her, and listening comprehension level, which is the level at which a person can comprehend material read aloud (Burns & Roe, 2002). The number of errors an individual makes in either word recognition or reading comprehension questions determines the reading levels. Typically, an IRI includes an assessment with graded sight-word lists that give a quick estimate of grade level. Once an estimate of an individual's independent grade level has been established, the individual reads passages beginning at their independent reading level and answers comprehension

questions about the passages (Burns & Roe, 2002; Paris & Carpenter, 2003; McCabe, Margolis & Barenbaum, 2001).

As informal assessments of reading, IRIs typically do not seek to establish psychometric properties to the breadth and depth found with standardized achievement measures (Johnson, Kress & Piluski, 1987; Nilsson, 2008; Paris, 2002; Paris & Carpenter, 2003; Spector, 2005). Despite this, IRIs attempt to establish validity and reliability for measures. For example, most IRIs are constructed using passages and words from a well established basal reader series and use a readability formula to determine reading difficulty (Burns & Roe, 2002; Fuchs, Fuchs & Deno, 1982; Gerke, 1980). In addition, empirical research has upheld the cut-off scores used to determine different grade levels. The reliability of IRIs has experienced slightly more controversy, with some indicating the reliability of most IRIs as adequate (Paris & Carpenter, 2003) and others criticizing the reliability of IRIs (Pikulski, 1974; Spector, 2005). Most IRIs simply do not report reliability figures, and few empirical studies have explored reliability. One study explored the reliability of three IRIs and found all three had alternate forms reliability of 0.60 to 0.78, suggesting moderate reliability (Helgren-Lempesis & Mangrum, 1986). In sum, little evidence exists to demonstrate the validity and reliability of IRIs, but what evidence does exist suggests moderate to good validity and reliability.

Academic skills in children with autism

Little research about the specific academic strengths and weakness of children with autism has been conducted. Some of the studies of the academic achievement of children with autism found unique patterns of deficits in academic achievement, such as strengths with word decoding but deficits in comprehension, but this finding is not consistent with all children with autism spectrum disorders (Church, Alisanski & Amanullah, 2000; Mayes & Calhoun, 2003a;

Mayes & Calhoun, 2003b; Mayes & Calhoun, 2008; Mineshew, Muenz, Goldstein, Payton, 1992; Minshew, Goldstein, Taylor & Siegel, 1994; Nation, Clarke, Wright & Williams, 2006).

With regards to reading, children with autism have demonstrated a significant range of skills (Griswold, Barnhill, Myles, Hagiwara & Simpson, 2002; Nation et al. 2006; Newman, Macomber, Naples, Babitz, Volkmar & Grigorenko, 2007). One pattern occasionally seen in children with autism is the demonstration of hyperlexia, typically defined as a deficit in reading comprehension despite having average abilities with word reading and decoding (Grigorenko, Klin, Pauls, Senft, Hooper & Volkmar, 2002; Grigorenko, Klin & Volkmar, 2003; Nation et al. 2006; Newman et al. 2007). In one study, researchers found 22% of their total sample of children 6- to 15- years of age were unable to read. Sixty-five percent of the remaining children demonstrated reading comprehension scores one standard deviation below population norms and 38% scored two standard deviations below population norms (Nation et al. 2006). In another study, twenty out of a sample of 41 individuals between 6- and 19-years of age who had been diagnosed with autism also demonstrated hyperlexia. For these children, sight word reading ability was similar to typically developing peers and greater than the sight word reading ability of children with autism, but the hyperlexic children struggled more than their typically developing peers in the areas of reading comprehension despite these similar sight word reading skills. Children with autism who were not diagnosed as hyperlexic demonstrated deficits with sight word reading and reading comprehension (Newman et al. 2007).

Differences in reading abilities also may be found based on diagnosis and the samples used within studies. For example, one study including 53 children with autism indicated children with an Intelligence Quotient (I.Q.) score of less than 80 had higher reading decoding skills than their I.Q. scores, but there were no differences found between reading and I.Q. among those

children with an I.Q. score of 80 or above. In another study, no differences were found between the Basic Reading and Reading Composite scores on the Wechsler Individual Achievement Tests (WIAT) for children with Asperger's Disorder, and these scores were all within in the Average range (Griswold et al. 2002). Similarly, another study including children who had been diagnosed with high-functioning autism found no significant differences between the Word Reading and Reading Comprehension composite scores on the Wechsler Individual Achievement Test – Second Edition (WIAT-II) and their scores were within the average range (Mayes & Calhoun, 2008). In another study, children with autism demonstrated average range reading skills on the WRAT-R (Szatmari, Tuff, Finlayson & Bartolucci, 1990). Based on these findings, it can be concluded no clear profile for the reading skills of children with autism and Asperger's Disorder has emerged in the literature.

Informal Reading Inventories have revealed children with Asperger's Disorder do not have significant differences between their instructional levels of reading and their actual grade level; however, their silent reading and independent reading levels were below their actual grade level. Children with Asperger's Disorder demonstrated greater success when answering literal/factual than inferential comprehension questions (Myles et al. 2002).

Children with autism also show no clear pattern with regards to written expression (Griswold et. al. 2002; Mayes & Calhoun, 2008; Mayes & Calhoun, 2003a; Mayes & Calhoun, 2003b). For example, Mayes and Calhoun (2008) found children with high-functioning autism demonstrated a lower mean score, 87, on the Written Expression subtest compared to their typically developing peers in the normative sample. However, another study examining children with Asperger's Disorder did not find a difference between their mean written expression scores and those of the children's typically developing peers within the standardization sample

(Griswold et al. 2002). Children with Asperger's Disorder may struggle more with the motor task of writing, rather than with the skill of expressing ideas through writing (Church et al. 2000; Myles et al. 2003). In one study, the children with Asperger's Syndrome tended to produce brief sentences, violations of rules of sentence construction (e.g., run-ons), and illegible letters at a significantly greater rate than their typically developing peers (Myles et al. 2003). In a qualitative chart review of 40 children diagnosed with Asperger's Syndrome, Church, Alisanski and Amanullah (2000) found written language, especially handwriting, was difficult for many of the children.

Mathematics has less often been described as an area of weakness of children with autism or Asperger's Disorder than reading and writing (Church et al. 2000; Mayes & Calhoun, 2008a; Mayes & Calhoun, 2003b; Griswold et al. 2002). In fact, qualitative reports of the abilities of children with Asperger's Syndrome indicated mathematics was an area in which many of the children excelled above their peers and their own achievement in other academic areas (Church et al. 2000). A comparison of the quantification and counting abilities of children with high-functioning autism revealed similar abilities as their typically developing peers (Gagnon, Mottron, Bherer & Joannette, 2004). Another study found Low Average range skills among children with autism as measured by the WRAT-R (Szatmari et al. 1990).

Goldstein, Minshew and Siegel (1994) suggested the academic performance of children with autism might vary along a developmental progression. Their study examined the academic performance of high functioning children (i.e. I.Q. scores above 70) diagnosed with autism on three standardized measures of academic achievement (Woodcock Reading Mastery Test, Revised [WRMT-R]; Woodcock, 1987, Kaufman Test of Educational Achievement [K-TEA]; Kaufman & Kaufman, 1985, Detroit Tests of Learning Aptitude-Second Edition [DTLA-2];

Hammill, 1985). Results indicated three typical patterns: some skills were average or above average regardless of age, other skills never achieved average levels in children with autism, and other skills began as average but were below average in older children, suggesting possible regression in these areas (Goldstein, Minshew & Siegel, 1994). It is important to note this study was not longitudinal, so progression of skills based on this data is hypothetical, rather than based on data.

One of the difficult components of assessing academic achievement in children with autism involves the use of the measures themselves. Due to the unique behavioral characteristics found in autism, completion of most standardized assessments requiring attention, effort, and engagement with the tasks may be difficult for many children with autism and therefore may not yield valid results. Indeed, children with autism have demonstrated significant difficulties completing Reading Comprehension subtests of academic achievement tests during empirical research (Mayes & Calhoun, 2003a; Mayes & Calhoun, 2003b). Children with autism who are older and those who have higher I.Q.s (e.g., above 80) may be more likely to successfully complete academic achievement tests (Mayes & Calhoun, 2003a; Mayes & Calhoun, 2003b).

Purpose of the study

Although summer camps have demonstrated success with treating behavior in children with autism (Hung & Thelander, 1978; Lopata et al. 2006; Lopata et al. 2008), more research is required to determine if academic skills can be improved through academic instruction over the summer months. In addition, the prevention of academic skill loss through the maintenance of academic skills in children with autism during the summer months also would be beneficial.

The purpose of the current study is to determine if academic achievement skills of children with autism can be maintained or improved over the course of summer camp with

academic programming. In particular, the WRAT4 (Wilkinson & Robertson, 2006) and an IRI (Burns & Roe, 2002) will be used to measure academic achievement during the first and last weeks of an eight weeklong summer camp. By exploring the use of a summer camp incorporating academic instruction to maintain or increase academic skills in children with autism, this study is intended to demonstrate a successful method of continuing educational services for children with autism during the summer months. This proposes to test one hypothesis and to answer one research question:

Research Hypothesis: Children with autism who participate in a summer camp including academic instruction will maintain or improve their reading, spelling, and mathematical computation skills as measured by the Wide Range Achievement Test – Fourth Edition.

Research question: Will children with autism who participate in a summer camp including academic instruction improve the functional, instructional, frustration reading levels and the listening comprehension level as measured by an Informal Reading Inventory?

Significance of the Study

This study will contribute to the literature regarding the use of therapeutic summer camps to deliver academic interventions for children with autism. Most studies utilizing a summer camp format to deliver interventions to children with autism have focused on behavioral, rather than academic, deficits (Lopata et al. 2006; Lopata et al. 2008; Hung & Thelander, 1978; Van Wert & Reitz, 1978). No studies were identified as seeking to increase the academic skills of children with autism. However, academic skills are necessary for independent

functioning (e.g., reading cooking instructions); thus, comprehensive treatment for children with autism should include an academic component.

The present study will contribute to the literature regarding the academic strengths and weakness of children with autism. There currently is a paucity of literature examining the academic strengths and weaknesses of children with autism. For example, one study has suggested children with Asperger's Disorder struggle with answering inference questions, compared to factual/literal questions on an IRI (Myles et al. 2002). By incorporating an IRI in this study, results of this study will add to the literature in this area. In addition, few studies have included a standardized measure of academic achievement with individuals with autism. Studies including a standardized measure of achievement indicated children with autism demonstrate a varied pattern of academic skills (Griswold et al. 2002; Mayes & Calhoun, 2008; Mayes & Calhoun, 2003a; Mayes & Calhoun, 2003b; Minshew et al. 1992; Minshew et al. 1994; Myles et al. 2003; Szatmari et al. 1990).

No studies examining the effectiveness of academic interventions for children with autism delivered over the course of the summer have been published to date. Therefore, this study will not only expand our knowledge of the unique patterns of academic skills demonstrated on standardized achievement measures by children with autism, it will illuminate the effects of academic instruction on those skills. Such findings may offer ideas for future researchers and practitioners who seek to provide such interventions for children with autism in their communities during the summer months.

Finally, this study will serve as a model for communities who wish to provide educational services for children with autism during the summer months, when public schools typically do not provide services. Summer camps are a logical option for communities wishing to

provide services for children with autism. Children with autism will likely find a summer camp appealing due to the active nature of most of the activities and opportunity to be outdoors.

Parents benefit from summer camps due to their relatively low cost compared to daycare services and the opportunity to have academic and behavioral interventions through the camp, which are often not available through daycare providers. In addition, counselors and teachers who provide services to children with autism will expand their knowledge of working with children with autism. Although there currently are examples of summer camps designed for children with autism (Lopata et al. 2006; Lopata et al. 2008; Hung & Thelander, 1978), this study will add another example to the currently existing literature available to communities which may wish to initiate their own summer camps.

CHAPTER II

Review of the literature

This chapter reviews the literature relevant to the current study. Specifically, research regarding the potential loss of academic skills during the summer as well as academically focused summer camps and programs are reviewed. Summer camps and programs for children with disabilities, which typically have a behavioral or psychosocial focus, are also reviewed. The academic achievement of children with autism is reviewed, and reviews of two academic assessments, the Wide Range Achievement Test (WRAT) and informal reading inventories (IRIs) are included.

Summer camps and programs for children with disabilities

The discussion about summer camps and programs for children with disabilities is complicated by the definition of the label “camp”, as no clear definition of what constituted a “camp” emerged in the literature. Summer programs have been conducted in a variety of settings ranging from, but not limited to, residential camping in a wilderness setting (Hung, 1977; Hung & Thelander, 1978), meeting daily in a wilderness setting (Eng & Davies, 1991), meeting daily on a university campus (Lopata, Thomeer, Volker & Nida, 2006; Lopata, Thomeer, Volker, Nida & Lee, 2008), and regularly including some adventure therapy techniques (e.g. ropes courses) into programs which otherwise did not include a wilderness setting (Jelalian, Mehlenbeck, Lloyd-Richardson, Birmaher & Wing, 2006). Some of these programs labeled themselves as “camps” while others as “programs”, and a clear correlation between label and setting, procedure or any other factor was not clearly found. For the purposes of this review, summer programs and camps without a strictly academic focus are included in this section, regardless of what setting or

procedures were used. Summer programs focusing solely on academic skills (summer schools) are considered separately.

In the past, many summer camps and programs included only typically developing individuals or only individuals with disabilities, but within the past few decades many camps have increasingly sought to integrate their programs (Blake, 1996; Bullock, Mahon & Welch, 1992; Goodwin & Staples, 2005). One of the arguments put forth by proponents for inclusion at summer camps that previously only included typically developing children was the concern that camps only for children with disabilities may not be able to provide services comparable to camps for children without disabilities (Goodwin & Staples, 2005). Other arguments include that camps designed specifically for children with disabilities may create sentiments of pity and they do not encourage the broader social justice goals of individuals with disabilities (Bullock et al. 1992; Goodwin & Staples, 2005).

However, it has also been suggested that inclusive camps for children with disabilities may also produce feelings of loneliness, isolation and disempowerment among campers with disabilities (Goodwin & Staples, 2005). Children who have participated in segregated camps have reported positive experiences, including that they felt less alone, felt encouraged by other campers to increase their independence, and had a chance to learn from other campers about the various possibilities available to children with disabilities (Goodwin & Staples, 2005). In addition, individuals with severe disabilities have been able to participate fully in traditional camp activities. In one example, eighty adolescents who had been institutionalized due to a variety of conditions (e.g., mental retardation, brain damage, schizophrenia) attended a segregated camp including traditional camp experiences, but no specific therapeutic interventions, rated the camp highly, and indicated they enjoyed the activities (Herr, 1977).

Summer camps have often distinguished themselves based on the type of disability, such as designing a camp solely for children with physical disabilities (Briery & Rabian, 1999; Eng & Davies, 1991; Jelalian et al. 2006; Klee, Greenleaf & Watkins, 1997; Martiniuk, 2003; Stevens, 2004). Other camps have been designed to focus solely on psychological disorders such as Learning Disabilities, Attention-Deficit Hyperactivity Disorder, and Developmental Disabilities (Bateman, 1968; Bennett, Cardone & Jarczyk, 1998; Davis-Berman & Berman, 1989; Eikeanaes, Gude & Hoffart, 2006; Forehand, Mulhern & Rickard, 1969; Gillis & Simpson, 1991; Heckel, Hursh & Hiers, 1977; Henley, 1999, Hinshaw, 2002; Hung & Thelander, 1978; Kelley, Coursey & Selby, 1997; Larson, 2007; Lattal, 1969; Romi & Kohan, 2004; Russell, 2003; Rynders, Schleien & Mustonen, 1990; Stich & Senior, 1984; Westervelt, Johnson, Westervelt & Murrill, 1998; Yssel, Margison, Cross & Merbler, 2005).

Camps and programs seeking to treat individuals with physical disabilities tend to focus on reducing the negative mental health impacts of the disorders (Briery & Rabian, 1999; Eng & Davies, 1991; Klee, Greenleaf & Watkins, 1997; Martiniuk, 2003; Packman, Fine, Chesterman, van Zutphen, Golan & Amylon, 2004; Powars & Brown, 1990) rather than addressing the physical disability directly (Jelalian et al. 2006). In one example, a one-week camp for children with serious illness was successful in making a significant reduction in the anxiety levels of campers regarding their conditions as measured by the Child Attitude Toward Illness Scale (CATIS). Another program used a ropes course in addition to traditional therapies to target morbid obesity in children and adolescents (Jelalian, et al. 2006). A comparison between those adolescents who completed the program, that also included a ropes course, revealed a statistically significant greater ability to maintain weight loss 10 months after treatment (Jelalian, et al. 2006). Another camp designed for children with physical disabilities (e.g., diabetes, cystic

fibrosis) focused on teaching children to manage their medical conditions and to make socio-emotional improvements by creating a normalizing experience with other campers with similar medical conditions (Shasby, Heuchert & Ganmeyer, 1984). Results indicated higher scores on the Piers-Harris Self-Concept Test (Piers-Harris, 1967) and the Nowicki-Strickland Locus of Control Scale for Children (Nowicki, & Strickland 1973), indicating the campers developed a more internalized locus of control and greater Self-Concept (Shasby et al. 1984). Similarly, a cross-country skiing program designed for adolescents who had hearing impairments successfully increased the internalized locus of control of participants as measured by the Levenson Locus of Control scale (Levenson, 1972), to a greater extent than a control group of age and hearing impairment matched peers (Luckner, 1989).

Several summer camps have reported success with integrating children with disabilities, and also indicated positive reactions from campers, staff, and parents. (Blas, 2007; Brookman et al, 2003; Bullock et al. 1992; Rynders et al. 1990). In one example of successful integration, a summer camp designed for Jewish children implemented a program, “Tikvah”, which was created in the 1970s and was designed to include children with a variety of disabilities (e.g., developmental disabilities, learning disabilities) in the general camp program. In another instance, a two-week summer day camp conducted in a traditional camp setting integrated three children with disabilities, one with autism and two with severe mental retardation, into a camp including eight typically developing peers (Rynders et al. 1990). In another camp, eight children with autism, with the assistance of personal aides, attended a camp designed for typically developing children. The aides used techniques such as priming, instruction on self-management skills, and facilitation of social interactions to help the campers with autism integrate into camp

activities (Brookman et al. 2003). Two of these camps (Blas, 2007; Brookman et al. 2003) used qualitative assessments and observations during camp and reported inclusion was successful.

Rynders et al. (1990) utilized behavioral data to examine the progress of campers. In this camp, the campers with disabilities demonstrated an increased ability to complete typical summer camp activities (e.g., meal preparation, swimming) as a result of being given social reinforcement. However, the campers actually showed a reduction in socially appropriate behaviors during crafts and outdoor education activities and no statistically significant increase in rates of social interaction (Rynders et al. 1990). Despite this, all the camps declared the integration of campers with disabilities was done successfully, though in each study this conclusion based on informal observations that the campers with disabilities were able to perform all camp activities and there was little to no attrition for campers with disabilities (Blas, 2007; Brookman et al. 2003; Rynders et al. 1990).

Although many camps have moved toward an integrated model, segregated summer camps designed only for children with disabilities continue to exist. One such camp designed for children with Attention-Deficit Hyperactivity Disorder (ADHD) delivered therapeutic interventions using typical camp activities (e.g., hiking). This five-week camp included 25 children ranging in age from 6- to 12-years-of-age. Counselors used techniques such as coaching, cueing, and redirection to complete tasks successfully. Formal assessment of the camp was not conducted, though the researchers cited parent reports indicating the campers had increased their abilities to control impulsive behaviors, increased their empathy with others, and reduce their aggression (Henley, 1999).

Four camps were found which specifically targeted children with Learning Disabilities (Michalski et al. 2003; Mishna, 2005; Westervelt et al. 1998; Yssel et al. 2005). All the camps

included an academic component as well as targeting the behavioral and emotional impacts of having a learning disability (Michalski et al. 2003; Westervelt, Johnson, Westervelt & Murrill, 1998). Two of the camps (Michalski et al. 2003; Westervelt et al. 1998) used standardized measures to measure the outcomes of the camps, and two (Mishna, 2005; Yssel et al. 2005) used only qualitative methods to evaluate their camps.

Yssel et al. (2005) conducted a weeklong summer camp for gifted but learning-disabled (“twice exceptional”) children aimed at increasing the campers’ self-esteem and confidence regarding academic tasks. The camp utilized educational tasks conducted in an outdoors setting. After the camp, parent feedback was obtained, and parents reported their children had made significant gains in self-esteem and confidence. Mishna (2005) utilized self-psychology techniques in a traditional summer camp setting for children with Learning disabilities. The premise of using self-psychology techniques was to emphasize empathy and understanding among campers and their counselors. Mishna (2005) reported two case examples of campers who, according to counselor reports, reduced the frequency of their behavior problems significantly during camp, though no data was given to support this assertion.

Michalski et al. (2003) reported statistically significant increases in self-esteem for children (13-years-old or younger) immediately after camp and at follow-up 6- to 8-months but adolescents (14-years or older) did not show statistically significant levels of improvement in self-esteem. Statistically significant reductions in feelings of loneliness, as measured by the Children’s Loneliness Questionnaire (CLQ) were found after camp for both children and adolescents and were maintained at follow-up. Self-report ratings of social competence, as measured by the self-report form of the Social Skills Rating System (SSRS) did not reveal significant changes from the beginning of camp to the end or at follow-up, but parent reports on

the SSRS did indicate significant improvement in social skills during camp and at follow-up. The authors indicated positive gains in overall self-concept as measured by the Self Description Questionnaire (SDQ) and increases in self-concept related to academic tasks. Interestingly, the authors reported children enrolled in camp who also had a diagnosis of ADD or ADHD did not experience as significant gains in self-concept as their peers (Westervelt et al. 1998).

Westervelt et al. (1998) also assessed the academic progress of the campers. Overall, campers demonstrated significant improvements in phonetic reading skills as measured by the Woodcock Johnson-Revised (WJ-R; Woodcock, 1989), phonetic reading skills, and spelling as measured by the Diagnostic Potential Spelling Test (DP; Arena, 1982) and word reading accuracy as measured by the Gray Oral Reading Test-Third Edition (GORT-3; Wiederholt & Bryant, 1992). Campers did not show significant gains with word recognition as measured by the DP or reading rate as measured by the GORT-3 (Westervelt et al. 1998).

In a program targeting adolescent substance abusers, a mixed methodology was used which included wilderness experiences (e.g., spending a week camping and completing a ropes course), and more traditional therapies, such as group therapy and a 12-step program. This program demonstrated a statistically significant decrease in staff reports of behavioral problems and self-ratings as measured by the Revised Behavior Problem Checklist (RBPC; Quay & Peterson, 1987). Participants also experienced a significant decrease on the Depressive (scale 2), Psychasthenia (scale 7), Schizophrenia (scale 8), and Hypomania (scale 9) clinical scales of the Minnesota Multiphasic Personality Inventory (MMPI; Hathaway & McKinnley, 1982) after completing the camp. Self-esteem of participants also increased significantly as measured by the Culture-Free Self-Esteem Inventory (CFSEI; Battle, 1981) (Gillis & Simpson, 1991).

One organization reported on four different day camps it had run for 70 children with severe mental retardation and reported mixed results (Bateman, 1968). These camps resulted in very slight increases in cognitive abilities, drawing abilities, and language abilities, but one of the camps saw a slight reduction in the skills of the campers who attended (Bateman, 1968). Procedural details regarding the activities or interventions used at the camp were not provided, so inferences about the different success rates among the camps was not possible. Baker (1972) also conducted a therapeutic camp for 25 children with mental retardation (mean I.Q. score of 41.7) using behavioral therapies (e.g., tokens, contingency management) to reduce a variety of behavioral problems and to increase adaptive behaviors. Behavior plans were individualized to each camper, and behavioral data collected used digits to indicate progress (e.g., -1 = worse, 1 = better) and numerical totals were then calculated to determine the overall progress of the camper. Results indicated campers showed more behavioral improvement than 15 matched controls that did not attend the camp (Baker, 1972).

McIntosh and Rawson (1988) conducted four summer camps, running for 10 weeks each, including a total of 114 children between 6 and 12 years of age. The camp utilized intensive behavioral remediation as well as group academic instruction and two individual tutoring sessions lasting 20 minutes each. Results of the camp indicated inconsistent results, but did suggest older children may have more success increasing internal locus of control through such a summer camp based on the Nowicki and Strickland Locus of Control Scale for Children (CNS-IE; Nowicki & Strickland, 1973). In a similar camp, Rawson and McIntosh (1991) conducted two summer camp programs for a total of 68 children with behavioral disorders. The camp sessions ran for 10 days each, with campers randomly assigned to camp sessions. Counselors and staff used a variety of behavioral therapy techniques similar to that of Rawson and McIntosh

(1991) to increase socially appropriate behaviors and self-esteem. Parent and self-report measures of self-esteem using the Culture-Free Self-Esteem Inventory (SEI; Battle, 1981) found significant increases in self-esteem in most areas (Rawson & McIntosh, 1991).

A camp specifically designed to treat children with emotional disturbances demonstrated effectiveness in reducing behavioral outbursts (crying and physical aggressiveness) of 2 campers by using a token economy system (Forehand, Mulhern & Rickard, 1969). Another summer camp program using contingency management increased compliance rates to 100% during the majority of camp days by all eight campers included in the program (Lattal, 1969). A similar camp targeting “disturbed” children (not defined further) found token reinforcement was effective at decreasing discussion of irrelevant environmental topics, increasing communication among members, and increasing problem solving and decision making skills (Heckel, Hursh & Hiers, 1977). Behavioral therapy in the form of token systems also was found to increase the academic productivity of children with emotional disabilities attending a therapeutic summer camp (Rickard, Clements & Willis, 1970; Rickards, Taylor & Libb, 1973). In one study, Rickard et al. (1970) offered remedial mathematics instruction and utilized contingency management skills to increase the amount of mathematics problems completed. The camp demonstrated a gain in mathematical skills as measured by the Iowa Test of Basic skills, and a loss of skills in the language portion, indicating the contingency management was successful at facilitating academic gains (Rickard et al. 1970). A similar study found the rate of responding on mathematics tests could be increased using contingency management (Rickard, Taylor & Libb, 1973).

Zwaig (1974) also conducted a summer camp using token reinforcement. Children enrolled in this camp were described as having learning disabilities, though behavioral descriptions of the campers indicated a wide variety of emotional dysfunctions. Results of this

program indicated a complete elimination of problem behaviors for all campers at the end of the two-week period, though no data were presented to support this claim (Zwaig, 1974). In another camp, children with learning disabilities or who were described as “educable mentally retarded” and who exhibited sensorimotor dysfunction were treated using a variety of activities designed to increase sensorimotor skills (e.g., horseback riding). Results of this camp indicated statistically significant increases in self-concept as measured by the Piers-Harris Children’s Self-Concept Scale (PHCSCS) (Zemke, Knuth & Chase, 1984).

In summary, a variety of summer camps for children with physical and psychiatric disabilities have been conducted using a variety of different methodologies. Whether inclusive or segregated, nearly all of the camps indicated positive results for campers, staff, and/or parents. Based on the literature, it appears clear summer camp programs can be used to effectively deliver therapeutic interventions for children with a variety of disabilities.

Academic summer programs

When formal schooling in the United States was in its infancy most communities were able to define their own schedules, allowing agricultural areas the ability to have summers off to allow children to participate in harvests on family farms and leaving urban schools free to have year-long school programs (Cooper et al. 1996). However, at the turn of the century schools sought to standardize curriculums used by schools, which led to nearly all schools adopting a calendar including a summer vacation (Cooper et al. 1996). For the past several decades, very few families have required their children to assist with farm operations, leaving them with a large span of free time during the summer. In response, many schools and community organizations began to create recreational summer programs for children, which over time came to be seen by parents and educators as a time to provide students with additional educational opportunities to

accelerate learning or provide academic remediation to students who struggled during the school year (Cooper et al. 2000).

Summer vacation may not only mean a cessation of academic instruction, but may actually result in a loss of skills during the summer (Cooper et al. 1996). In one meta-analysis of 13 studies examining learning loss in the summer, Cooper et al. (1996) found students at best made no academic gains during the summer and at worst lost one tenth of a standard deviation on achievement tests between spring and summer assessments. In addition, this estimate may be an underestimate of loss, as students in these studies were not typically tested on the very last and first days of school, indicating some instructional time between testing likely occurred and may have mitigated some loss of skills (Cooper et al. 1996).

Many studies have reported children from low socio-economic status (SES) households may be particularly likely to experience achievement losses during the summer. In one study that followed the academic progress of 6th and 7th grade students in the Atlanta public schools for two years, including one summer session, Heynes (1978) found the loss of achievement skills during the summer was significantly more pronounced for children from low SES households. This study did not show a significant loss of achievement skills among students once the SES of students was taken into account (Heynes, 1978). In another investigation, Kim (2004) followed over 1,600 elementary students from a public school district and found reading over the course of the summer lessened the loss of academic skills during the summer regardless of previous achievement level. In addition, the number of books read was positively correlated with SES (Kim, 2004), further indicating children with lower SES often have less access to reading materials. A third investigation following 787 elementary school students for 5 years found students from low SES households lost almost 2 points on standardized assessments over the

course of the summer, where their high SES peers made gains of nearly 47 points (Alexander, Entwisle & Olson 2007). These investigations all indicate the importance of SES as a moderator of achievement during summer vacation. Overall, results of these studies reinforce the finding of Cooper et al. (1996) of the potential for summer vacations to result in a loss of skills, and suggest the loss may be particularly significant for children from low SES backgrounds.

Summer programs targeting children who are from low SES backgrounds have indicated effectiveness in producing academic gains (Schacter, 2003). Schacter (2003) conducted an 8-week summer reading program that targeted 61 children from low SES backgrounds. This camp utilized a variety of typical camp activities such as swimming and putting on plays, but also included two hours of direct reading instruction using the SRA Open Court Reading series (McGraw Hill, 2000) per day and each child also had two 45-minute sessions per week with an individual tutor. Standardized achievement measures (the Gates-MacGinitie Vocabulary Level 1, fourth edition and the Gates-MacGinitie Comprehension Level 1, fourth edition [Gates-MacGinitie Reading Tests, 2000], the CORE phonics survey [Honing, Diamond & Nathan, 2000] and the Fry Oral Reading Survey [Fry, 1990]) indicated significant improvements in understanding of vocabulary, reading comprehension, and phonics skills, but did not indicate significant changes in oral reading (Schacter, 2003). In a similar program, Schacter and Jo, (2005) conducted a 7-week camp using the same curriculum as Schacter (2003), but used slightly different outcome measures (Gates-MacGinitie Word Decoding Levels 1 and 2 Form S [Gates-MacGinitie Reading Tests, 2000], Gates-MacGinitie Comprehension Level 1 and 2 Form S [Gates-MacGinitie Reading Tests, 2000], and the Stanford 9 Decoding and Comprehension tests Primary 2, Form T [Harcourt Brace Educational Measurement, 1997]). Instruction for the Schacter and Jo (2005) camp also was very similar as the Schacter (2003) camp, but did not

utilize individual tutoring. Schacter and Jo (2005) also included 72 disadvantaged students in their camp and utilized a control group of 90 matched peers. Results indicated that children in the reading camp group demonstrated significant gains compared to controls with comprehension skills immediately after camp and that these gains were maintained at three month as well as one year follow ups. Results also indicated children in the camp experienced improvements with decoding compared to controls immediately after camp and at three month follow up, but demonstrated similar skills as the control group after one year (Schacter & Jo, 2005).

Little research exists about academic summer camps or programs for typically developing or students with average academic achievement. One probable reason for this may be the cost of running such a program. It is understandable cash-strapped schools seeking to increase the achievement of students would likely focus on students who are under-achieving. As a result, most programs tend to target specific populations rather than accelerating the learning of children with average academic achievement.

In their review of seven studies focusing on the effects of summer programs increasing the achievement of gifted students, Cooper et al. (2000) determined the programs had an average effect of increasing achievement by about one sixth of a standard deviation. In another study, a summer camp format was used to provide additional mathematics instruction to 120 inner city African-American youth in middle and high school (Edwards, Kahn & Brenton, 2001). Results after six years of program implementation indicated significant gains overall, with students earning between 28 points and 60 more points on posttest measures of achievement. It is important to note these tests were not standardized and were created by the instructors of the camp (Edwards et al. 2001), and this may have influenced the results.

Summer programs also have been suggested as a means for successfully remediating students with learning disabilities. One investigation of a six-week, individualized summer school program for children with learning disabilities found significant increases in mathematics, though this increase was not maintained six-months after the completion of the program (Leviton & Kiraly, 1975). However, Leviton and Kiraly (1975) do not report significant detail to evaluate the program completely. In addition, no information about the instruction the students received after the completion of the program was provided, and it is possible the students' failure to maintain achievement was reflective of poor instruction after the program. In a more complete investigation, Cooper et al. (2000) conducted a meta-analysis of summer school programs with primarily an academic focus. Included in this study was a separate analysis of 41 programs targeted at individuals with "learning difficulties", described by various programs as "at-risk", "underachievers or below grade level", "failing or retained" and "learning disabled or otherwise impaired". These programs reported, on average, an increase of about one fifth of a standard deviation greater on outcome measures than control groups did and the average effect size for these programs was $d = 0.26$. In addition, results indicated students labeled learning disabled or having physical or mental defects reported the highest effect sizes of all "learning deficient" students (Cooper et al. 2000). Overall, these studies suggest children with learning disabilities may benefit from summer remediation programs, but suggest attention should be paid to long-term outcomes and instruction seeking to maintain gains made in summer programs.

Although this effect size is considered "small" according to Cohen's guidelines, the importance of academic skills to daily functioning, and to the pursuit of higher education and higher paying jobs, means even a small improvement of these skills is a valuable outcome. Further, given the serious deficiencies some of the children included in these studies, such as

those with learning disabilities, significant changes over the course of only a few weeks or month is not theoretically likely, making a even a small improvement an impressive feat. Nonetheless, it is important to note all studies utilized different methodologies, and therefore, individual program may exhibit varying effectiveness.

Academic achievement of individuals with autism

There is a paucity of literature examining the reading, writing, and mathematics skills of children with autism. The behavioral and communication deficits central to the diagnosis of autism may explain this lack of literature. Achievement measures, whether formal or informal, necessarily require interaction with an examiner and engagement to the tasks. However, social withdrawal, a lack of verbal skills and difficulty in communication are hallmarks of autism (APA, 2000). Therefore, it is understandable researchers have found it difficult to assess these skills within children with autism. In addition, there is a great deal of variation in the behaviors exhibited by children who have been diagnosed with autism. For example, children with high functioning autism may closely resemble their peers with Asperger's Disorder with the exception of a history of language delays, and therefore research attempting to find differences between these two groups may fail to find significant results.

Reading

Many researchers have described the reading scores of children with autism spectrum disorders as being greater than their peers, particularly with mechanical aspects of reading, such as reading individual words, but lower on tasks which require using inference, contextual clues, and comprehension reading skills (Church et al. 2000; Frith & Snowling, 1983; Minshew et al. 1994; Whitehouse & Harris, 1984; Rumsey & Hamburger, 1990; Mayes & Calhoun, 2003b; Nation et al. 2006). Consistent with this finding, some have asserted that children with autism are

more likely than their typically developed peers to exhibit hyperlexia (Whitehouse & Harris, 1984). For example, a retrospective chart review of 20 children with autism revealed a similar pattern of greater word reading skills on the Word Reading subtest of the WRAT than comprehension skills as measured by the Gates-McGinitie Reading Tests (1964) in most of the children studied, although not all met criteria for hyperlexia (Whitehouse & Harris, 1984).

Other researchers compared the abilities of children with autism to read individual words in a variety of contexts (Frith & Snowling, 1983). In this study, children with autism were compared to typically developing peers and dyslexic peers. The children with autism were not found to struggle compared to typically developing peers with phonetics or on a Stroop color-word interference task, but did differ significantly with pronouncing homographs and correctly completing sentences with missing words, suggesting they struggle to integrate contexts of sentences when reading (Frith & Snowling, 1983).

In another study, Nation et al. (2006) compared scores obtained by a group of 41 children between 6 and 15 years of age with diagnoses of autism, “atypical autism” and Asperger’s Disorder on The Graded Nonword Reading Test (Snowling, Stothard & McLean, 1996), the British Ability Scales (BAS-II; Elliot, Smith & McCulloch, 1996) and the Neale Analysis of Reading Ability-II (NARA-II; Neale, 1997). Results from these assessments indicated approximately 35% of the participants had word reading skills a full standard deviation greater than reading comprehension skills (Nation et al. 2006). An examination of the reading skills of children with autism using an earlier version of the same assessment, the Neale Analysis of Reading Ability (Neale, 1966) reported similar findings (Rutter & Bartak, 1973). In this study, 13 children out of a sample of 45 ten year old children with autism had reading accuracy rates

similar to or greater than their eight year olds, but only two children demonstrated reading comprehension skills of eight years or older (Rutter & Bartak, 1973).

Researchers in another study (Minshew et al. 1994) compared the scores of 54 individuals with autism, who had overall cognitive abilities above 70, on the reading subtest of the Detroit Tests of Learning Aptitude-2 (DTLA-2; Hammill, 1985), the Woodcock Reading Mastery Tests-Revised (WRMT-R; Woodcock, 1987), and the Kaufman Test of Educational Achievement (K-TEA; Kaufman & Kaufman, 1985). Results indicated the children with autism performed similarly to their peers on concrete tasks (e.g., Word Comprehension on the WRMT-R) but struggled compared to their peers on tasks involving comprehension (e.g., Reading Comprehension on the K-TEA) (Minshew et al. 1994).

In another chart review of 164 children with autism presenting to a psychiatric clinic and who completed the Wechsler Individual Achievement Test (WIAT) and the Woodcock-Johnson Tests of Achievement-Revised (WJ-R) indicated school-aged children with cognitive abilities lower than 80 had significantly greater reading decoding skills than their overall cognitive abilities (Mayes & Calhoun, 2003a). In a similar study, a chart review of 53 children using the same assessments revealed children with autism who had overall cognitive abilities of greater than 80 did not have significant differences between reading and overall cognitive abilities. Children with overall cognitive abilities of 80 or less demonstrated greater reading scores than overall cognitive abilities, and these children also struggled to complete tests of reading comprehension (Mayes & Calhoun, 2003b).

Also consistent with the above findings, Mayes and Calhoun (2008) found no deficits in reading comprehension compared to word reading abilities as measured by the WIAT-II in children with autism who had overall cognitive abilities of 70 or greater. A different assessment

instrument, WRAT-R (Jastak & Jastak, 1984), measured word reading abilities among 43 children with Asperger's Disorder and Autism, with overall cognitive abilities greater than 65, indicated greater word reading skills than overall cognitive abilities, and this pattern was not found among matched controls who had other psychiatric disorders (e.g., ADHD, Conduct Disorder) (Szatmari, Tuff, Finlayson & Bartolucci, 1990). A review of the school files of 41 children diagnosed with high functioning autism, pervasive developmental disorder, or PDD-NOS found mean overall reading skills within the average range ($M = 92.00$) on the Woodcock-Johnson Psycho-Educational Battery (Woodcock & Johnson, 1977). Subtest scores were not provided, so a comparison of word reading and reading comprehension skills was not possible (Myles, et al. 1994).

Studies including only children with Asperger's Disorder have not shown deficits in reading composite scores compared to mathematics, language, and written language composites as measured by the WIAT (Psychological Corp, 1992) (Griswold et al. 2002). The functional reading levels of children with Asperger's Disorder were examined in one study using the Classroom Reading Inventory (CRI; Silvaroli, 1993). The results of this study found no statistical differences between instructional reading level and grade level of the participants (Myles et al. 2002). This suggested children with Asperger's Disorder did not demonstrate significantly lower overall reading capabilities than their typically developed peers. However, this study did find a significantly greater ability among the children with Asperger's Disorder to answer reading comprehension questions which were factual/literal in nature rather than inferential (Myles et al. 2002). These findings suggest that although children with Asperger's Disorder may not generally have deficits in reading comprehension, they may nonetheless struggle compared to their typically developing peers at making inferences from reading.

Goldstien et al. (1994) examined achievement scores of 64 children with autism according to age. This study found children with autism tended to consistently perform better than controls on some tests, such as phonetic tasks, at all ages. Other skills, such as those involving complex linguistic material, did not equal those of their peers at any age. Other skills, such as reading comprehension, for children with autism were similar to those of their typically developed peers at younger ages (Goldstein et al 1994). However, older children with autism struggled compared to their peers on these tasks (Goldstein et al. 1994), suggesting a strong early skill set in this area but less growth in this area as their peers.

The results of the discussed studies suggest there is no clear pattern of reading skills for all children with autism spectrum disorders. For example, children with autism may demonstrate a pattern of greater abilities with word reading rather than reading comprehension, but the presence of this pattern of skills may vary based on factors such as cognitive abilities, diagnosis, and age.

Written Language

Regarding writing skills, some evidence exists to suggest individuals with autism are more likely to demonstrate deficits with written expression than their overall cognitive abilities (Mayes and Calhoun, 2003a; Mayes and Calhoun, 2003b; Mayes and Calhoun, 2008). For example, in one study, children with autism who had overall cognitive abilities greater than or equal to 80 demonstrated significantly poorer performance on written expression and spelling subtests than their overall cognitive abilities (Mayes & Calhoun, 2003a). This pattern of deficits in written expression abilities also was found for a different study including 54 children with autism who had overall cognitive abilities of 70 above (Mayes & Calhoun, 2008). However, in another study, children with autism who had cognitive abilities of 80 or greater demonstrated

lower written expression skills as measured by the WIAT, but those tested with the WJ-R did not reveal significant difference on the writing composite compared to other academic skills (Mayes & Calhoun, 2003b). These findings suggest children with autism struggle more than their peers with written expression, though these findings may be mitigated by overall cognitive functioning, as well as the assessment used to measure written expression skills.

In one study of 21 children and adolescents with Asperger's Disorder, average scores on the WIAT writing composite ($M = 95.82$) did not appear to be significantly different than the composite scores by the participants in reading ($M = 101.70$), Mathematics ($M = 91.76$) or Language ($M = 101.48$), but due to one participant refusing to complete the written expression subtest, statistical comparisons were not performed (Griswold, 2002). In another study, 16 individuals with Asperger's Disorder were administered the Test of Written Language (TOWL-3; Hammill & Larsen, 1996) and were found to have similar written expression skills as their typically developing peers. However, further analysis revealed children with Asperger's wrote significantly fewer words per sentence compared to their typically peers. In addition, analysis of the handwriting of the children with Asperger's Disorder using the Evaluation Tool of Children's Handwriting (ETCH; Amundson, 1995) revealed the majority of the children with Asperger's Disorder had significantly poorer handwriting compared to their typically developing peers (Myles et al. 2003). A qualitative analysis of educational records for children with Asperger's Disorder revealed similar findings, indicating many school-aged children with Asperger's Disorder struggled with handwriting and written expression within the classroom (Church et al. 2000).

Mathematics

Research regarding the mathematical skills of children with autism has generally suggested no differences in mathematical skills among children with autism and their typically developing peers (Church et al 2000; Goldstein et al. 1994; Griswold, 2002; Minshew, 1994; Minshew 1992) and overall cognitive abilities (Mayes & Calhoun, 2003a; Mayes & Calhoun, 2003b). In addition, some researchers have hypothesized that the unique deficits in processing which favor literal and factual learning in autism may explain the higher functioning of children with autism in mathematics and also may suggest children with autism would show superior skills compared to their typically developing peers in the areas of mathematics (Gagnon et al. 2004). However, studies examining the quantification and counting abilities of individuals between the ages of 10 and 21 years of age did not reveal greater skills than their typically developing peers (Gagnon et al. 2004). A qualitative review of educational records for 41 children with high functioning autism, pervasive developmental disorder, or PDD-NOS revealed a mean overall mathematics score in the low average range ($M = 80.00$) (Myles et al 1994). Based on these findings, it appears clear children with autism, in general, show greater skills in the area of mathematics than reading and writing, though like these other skills, there exists significant variability.

In summary, children with autism show a great amount of variability in functionality (Myles et al 1994) and their academic achievement reflects this heterogeneity. With regards to reading, children with autism may have a tendency to demonstrate weaknesses with reading comprehension even when word decoding and word identification skills are intact. In some children with autism, this discrepancy may be severe enough to warrant a diagnosis of hyperlexia, though not all children who exhibit this pattern will meet such criteria. Similarly,

children with autism may demonstrate deficits with written language, though there exists a significant amount of variability in this area as well. In general, children with autism demonstrate average range mathematical skills, but do not necessarily have a strength compared to typically developing peers in this area.

Wide Range Achievement Test

Description. The Wide Range Achievement Test, Fourth Edition (WRAT4; Wilkinson & Robertson, 2006) is an achievement assessment consisting of tests in reading, math, and spelling. Predecessors of the WRAT4 have historically been popular assessments of achievement skills (Piotrowski & Keller, 1989; Snelbaker, Wilkinson, Robertson, & Glutting, 2001; Sullivan & Hawkins, 1995). No doubt one of the contributing factors of the popularity of the WRAT4 and its predecessors has been the quick administration time, as administration of the WRAT4 can be done in 15 to 45 minutes (Wilkinson & Robertson, 2006). Indeed, the authors of the WRAT4 specify the focus of the WRAT4 is to provide a “quick, simple, psychometrically sound assessment of important fundamental academic skills” (Wilkinson & Robertson, 2006).

The WRAT4 consists of four tests: Word Reading, Sentence Comprehension, Spelling, and Math Computation. The Word Reading test involves letter identification and word recognition. The Sentence Comprehension subtest uses a modified cloze format in which a reader must suggest a word missing from a sentence. The Spelling test measures an individual’s skill with spelling individual letters and words. The Math Computation test measures an individual’s skill with completing paper and pencil mathematical calculations. The tests on the WRAT4 may be administered in any order, though the standardization sample completed the tests in the following order: Word Reading, Sentence Comprehension, Spelling, and Math Computation. The WRAT4 is intended to be an individualized assessment; however, the Spelling

and Math Computation tests may be administered in groups for individuals 8 years or older (Wilkinson & Robertson, 2006).

History. The WRAT was first published in 1946 and was meant to augment the Wechsler-Bellview scales of intelligence measures (Wilkinson & Robertson, 2006). The WRAT then underwent several minor changes before the WRAT-R was published in 1984 (Spruill & Beck, 1978; Vance, Kitson & Singer, 1985). The WRAT-R was created after two major changes to the original WRAT: the use of Rasch scaling of items and a significantly updated normative sample (Reid, 1986; Witt, 1986). The third edition of the WRAT, the WRAT3, was published in 1993 and did not reveal significant changes from the WRAT-R (Snelbaker, Wilkinson, Robertson, & Glutting, 2001). Clearly, a pattern can be seen in which each edition of the WRAT has drawn heavily on its most recent edition for item content. The lack of radical changes to the item content of the WRAT is likely due to the focus of the assessment, as core achievement skills in mathematics, spelling, and reading have not undergone significant changes since the first publication of the original WRAT (Snelbaker et al. 2001).

The WRAT4 was published in 2006 and drew heavily on its most recent predecessor, the WRAT3, for content and format. The WRAT4, like the WRAT-3 includes three forms: Blue, Tan, and Combined. The WRAT3 contained norms for individuals ranging in age from 5 to 74 years and were divided into 23 age groups. Age ranges were smaller for younger children and increased in size for older children and adults. The standardization sample included 4,433 individuals stratified according to the 1990 U. S. census data. The WRAT4 includes the Spelling, Word Reading, and Math Computation tests found on the WRAT3 with only minor revisions. The Sentence Completion subtest is new to the WRAT4 and was included at the request of users

of the WRAT3 who desired a means to measure reading comprehension (Wilkinson & Robertson, 2006).

Psychometric properties. Research studying the psychometric properties of the WRAT4 is somewhat limited, though this likely is due to the recent publication of the WRAT4 (Dell, Harold & Dell, 2008). However, the WRAT4 shares content and structure with the WRAT3, which has a significant amount of research demonstrating the efficacy of its psychometric properties. Because each edition of the WRAT has drawn heavily on previous versions for all revisions, the consideration of the psychometric properties of previous versions of the WRAT is warranted.

Validity. The WRAT demonstrated good concurrent validity with the Wechsler Intelligence Scale for Children – Revised (WISC-R; Wechsler, 1974) (Brooks, 1977; Naglieri, 1980; Reynolds, Dappen & Wright, 1981; Sattler, & Ryan, 1981; Schwarting & Schwarting, 1977), the Wechsler Intelligence Scales for Children (WISC; Wechsler, 1949) (Brooks, 1977), the Stanford-Binet Intelligence Scale (SB; Terman & Merrill, 1973) (Brooks, 1977), as well as with the (WIPPSI; Wechsler 1967) (Reynolds et al. 1981), and Kaufman Assessment Battery for Children (K-ABC; Kaufman & Kaufman, 1983) (Karnes, Edwards, & McCallum, 1986). The WRAT also has been found to correlate well with several achievement measures, such as the Peabody Picture Vocabulary Test – Revised (PPVT-R; Dunn & Dunn, 1981), and the McCarthy screening test (McCarthy, 1978) (Vance, Hankins & Brown, 1986) and the California Achievement Test (Karnes et al. 1986).

The content and format of the WRAT-R did not offer significant updates to the WRAT (Reid, 1986). On one comparison of the 1978 WRAT and the 1984 WRAT-R, the WRAT-R was found to produce standard scores that ranged from 8 points lower to 13 points lower when

compared to the WRAT, though this difference was attributed to the updated normative sample used for the WRAT-R (Stevenson, 1990; Spruill & Beck, 1986b). One criticism of the WRAT-R was the shortened length of the Spelling test, which reduced reliability and validity (Sullivan & Hawkins, 1995). The WRAT-R has been correlated with the Wechsler Adult Intelligence Scale-Revised (WAIS-R; Wechsler 1981) and the WISC-R (Cooper & Fraboni, 1988; Smith, Smith, & Dobbs, 1991; Spruill & Beck, 1986a). Measures of concurrent validity with the Kaufman Tests of Educational Achievement (K-TEA; Kaufman & Kaufman, 1985) have suggested the WRAT-R maybe produce standard scores between 4 and 10 points lower than the K-TEA (Prewett, Bardos, & Fowler, 1990; Prewett & Giannuli, 1991; Webster, Hewett, & Crumbacker, 1989).

The next version of the WRAT, the WRAT3, also demonstrated moderate concurrent validity between full scale I.Q. scores on the Wechsler Intelligence Scale for Children – Third Edition (WISC-III; Wechsler, 1991), with researchers obtaining correlations of 0.50 to 0.82 (Smith, Smith, & Smithson, 1995; Vance & Fuller, 1995). The WRAT3 has demonstrated correlations of 0.74 and 0.80 with the Kaufman Brief Intelligence Test (K-BIT; Kaufman & Kaufman, 1990) and 0.81 and 0.87 with the Kaufman Functional Academic Skills Test (K-FAST; Kaufman & Kaufman, 1994) (Klimczak, Bradford, Burright, & Donovanick, 2000).

The ability of an achievement test to reflect the functional achievement of children in the classroom also is of importance. The WRAT has been criticized as not having a rigorous enough curriculum for the identification of learning disabilities for younger children, as a six-year-old child would have obtained a standard score of 101, within the average range, if she or he could name 13 letters, recognize 10 letters, and name two letters in his or her name (Grossman, 1981). However, a comparison between school placement and the more current WRAT3 indicated results on the WRAT3 predicted the educational placement of 72% of children with learning

disabilities, 83% of children with educable mental handicaps and 85% of children receiving gifted program (Snelbaker et al. 2001). The differences in these results likely reflect the changes in normative samples between the WRAT and WRAT3.

Empirical studies of the validity and reliability beyond those presented in the technical manual of the WRAT4 are not currently available, likely due to the recent publication of this edition. However, the psychometric evidence related to previous editions of the WRAT and the similarity in content between the editions suggests the validity of the WRAT4 is likely to be good. Indeed, the WRAT4 manual cites validity studies for previous editions as part of the evidence for the validity of the current edition. It should be noted that the WRAT4 contains an entirely new subtest, for which psychometric studies of previous versions are not available. However, the manual presents evidence of validity which indicates the Sentence Comprehension subtest has sufficient external and internal validity (Wilkinson & Robertson, 2006). One independent review of the WRAT4 proclaimed that it demonstrates acceptable levels of content, internal, and external validity (Dell, Harold & Dell, 2008).

Reliability. Research on the reliability of the WRAT has not been as plentiful. A review of the WRAT3 indicated a correlation of 0.98 between the Blue and Tan form and coefficient alphas comparing each of the nine WRAT3 tests were between 0.85 and 0.95 and 0.92 to 0.95 for the combined form (Knoop, 2004; Snelbaker et al. 2001). Rasch person separation indices indicated the range of person separation was 0.98 and 0.99 on the nine WRAT3 tests (Snelbaker et al. 2001). Test-retest reliability was found to range from 0.91 to 0.98 on the nine test of the WRAT3 (Snelbaker et al. 2001).

Knoop (2004) reported standard errors of measurements (SEMs) for standardized scores on the combined form of the WRAT3 ranged from 10 to 12 points for 95% confidence intervals,

which may be larger than ideal. However, Snelbaker et al. (2001) report SEMs ranged from 3.4 to 5.9 on all forms, suggesting some disagreement in this area. Knoop (2004) suggests the larger SEMs indicate the WRAT3 should be used as a screening test only. This suggestion is commensurate with the WRAT4 manual, which states the purpose of the WRAT4 is to be a short, but psychometrically sound, assessment of achievement skills (Wilkinson & Robertson, 2006).

The WRAT4 also demonstrates good test-retest reliability. The WRAT4 manual reports immediate test retest reliability for all ages on the combined form of 0.88 and a delayed coefficient of 0.84 (Wilkinson & Robertson, 2006). Some have suggested test administrators use caution when using the WRAT4 for retesting (Dell et al. 2008). However, these coefficients are still quite high and this conclusion appears overly conservative. Although admittedly sparse compared to the evidence of validity for the WRAT, the evidence that does exist on the WRAT4 suggests adequate reliability.

All versions of the WRAT have been used successfully with special populations. The WRAT has been frequently used with children who have been referred for psychological evaluation (Brooks et al. 1977; Naglieri, 1980; Schwarting & Schwarting, 1977; Spruill & Beck, 1986). The WRAT-R reading subtest was used with 39 individuals with a variety of cognitive dysfunctions (e.g. traumatic brain injury [TBI], cerebral vascular accident [CVA]) and was found to be a stable estimate of premorbid intelligence for those in the sample who were experiencing cognitive decline over the course of approximately two years (Johnstone & Wilhelm, 1996). Similarly, the test-retest reliability of the reading subtest of the WRAT-3 with geriatric populations ($M = 71.7$ years-of-age) consisting of individuals with and without dementia revealed a test retest reliability of 0.90 across an average of 14 months. The authors of the

WRAT4 included several children with disabilities in the standardization sample (Dell et al. 2008; Wilkinson & Robertson, 2006).

In summary, the WRAT has been a popular tool for assessing achievement skills in children and adults since the publication of its first edition in 1946 (Snelbaker et al. 2001; Wilkinson & Robertson, 2006). Each subsequent version of the WRAT has drawn heavily on its predecessor for content and form, with the only major changes occurring with an update to the normative sample in between the original WRAT and the WRAT-R, and the addition of the Sentence Comprehension test on the current WRAT4 (Wilkinson & Robertson, 2006). The psychometric properties of the WRAT3, the most recent predecessor to the WRAT4, appear to be adequate, and are commensurate with the current WRAT4's stated purpose as concise but accurate measure of achievement abilities.

Informal Reading Inventories

Description. An IRI typically determines the reading level of an individual by measuring his or her skills when reading individual words and when comprehending written passages. The reader begins by reading aloud a graded word list, and his or her success with reading these words provides a rough estimate of his/her reading level. After this estimate is obtained, the reader begins to read passages, usually aloud, corresponding to the reader's independent reading level, which corresponds to the last word list in which they made no errors. While the child reads, the examiner follows along on a protocol, tracking miscues (omissions, substitutions, etc). The examiner may then analyze the miscues at a later time to determine if the child is making specific, repeated miscues which may be addressed through instruction. After the reader has finished reading the passage, he or she answers questions about the passage to assess the reader's level of reading comprehension.

Reading levels are typically determined by the number of errors made while reading aloud and by the number of comprehension questions missed. Reading levels are divided into Independent, Instructional, and Frustration levels. In addition, many IRIs include a Listening Comprehension section. The Independent level indicates a level at which a reader can read with very few errors and demonstrate comprehension through answering several comprehension questions correctly. For example, 1% or fewer errors while reading aloud and correct answers to 90% or more comprehension questions (Burns & Roe, 2002). The Instructional level indicates a level in which a reader can read fluently and comprehend text with the assistance of a teacher. For example, 85% of all words read aloud correctly and answering 75% of comprehension questions correctly (Burns & Roe, 2002). The frustration level is a level of reading in which, even with the help of a teacher, fluency and comprehension are not likely to be obtained. For example, reading fewer than 85% words correctly and answering less than 50% of comprehension questions correctly (Burns & Roe, 2002). Listening comprehension corresponds to what material a student can comprehend when it is read aloud to him or her. For example, the level at which an individual can correctly answer 75% or more of comprehension questions regarding a passage read aloud to them (Burn & Roe, 2002).

The reading passages included in IRIs are usually selected with the goal of representing paragraphs that could be found in reading materials and textbooks corresponding to a given grade level. The most common resource of material for constructing IRI are passages from graded basal readers or textbooks which can be found in classrooms (Brecht, 1977; Burns & Roe, 2002). An IRI may be constructed by a classroom teacher from his or her own classroom materials or one of the many commercially produced IRIs available may be used (Pikulski & Shnahan, 1982).

Unlike a standardized reading assessment, an IRI focuses on determining the functional level of an individual's reading skills rather than specific reading skills (Kaufman & Kaufman, 2004; Wechsler, 2001; Woodcock-Johnson, 2001). Standardized achievement measures offer a means to compare multiple aspects of an individual's reading skills with those of his or her peers. However, the tasks included on standardized measures have been criticized as not correlating with an individual's functional reading (Berger & Kautz, 1967; Fuchs, Fuchs, & Deno, 1982; Lovitt & Hansen, 1976). For example, a child may be able to read vocabulary commensurate with his or her peers, but he or she may struggle to comprehend material presented in textbooks or novels. There is some limited empirical evidence to suggest that standardized achievement tests may not accurately assess a reader's functional reading level (Berger & Kautz, 1967).

Teachers may prefer IRI assessments to standardized assessments due to greater convenience and face validity (Davis, 1975; Gonzales & Elijah, 1975). Indeed, research has suggested a majority of teachers have an awareness of IRIs and have used a variety of different forms of IRIs in their classrooms (Masztal & Smith, 1984). Teachers may find IRIs more appealing because they believe IRIs offer more functional information about reading compared to standardized achievement measures (Burns and Roe, 2002). The measurement of different levels of reading offer more utility to a classroom teacher than a standardized achievement score when placing a reader in an instructional program and when selecting reading materials that are appropriately challenging to the individual. In addition, many IRIs offer a miscue analysis component for passages read aloud which can offer insight into the types of reading errors a student may be making (Nilsson, 2008; Johnson, Kress, & Pikulski, 1987).

History. The beginnings of the IRI are not as well documented as the beginnings of other assessments, such as standardized achievement measures (Davis & Eckwall, 1976). Educational professionals, such as classroom teachers, have long required methods to assess the functional reading performance of a student in order to ensure the student is provided with instructional methods and materials commensurate with his or her skill level in order to maximize learning. Teachers have long used a format similar to the IRI to determine grade placement for reading. A relatively straight forward and logical method has been to have a child read materials available within the classroom or school system and monitor his or her success rate to determine where a child was reading functionally (Anderson, 1977). The development of written guides to aid teachers in the construction of IRIs and the development of commercial IRIs has sought to increase the strength and utility of this informal method.

Johns (1983) traces the history of the IRI back to Waldo (1915), a superintendent of the Public Schools of Illinois as being one of the first professionals to use an assessment similar to current IRIs. In particular, John (1983) credits Waldo (1915) as a pioneer in his measurement of silent reading rate and comprehension. However, he notes the primary function of this evaluation was more concerned with judging teacher effectiveness and improving teacher's instructional methods than assessing individual students (Johns, 1983). At approximately the same time, Gray developed assessments of reading which also included silent and oral reading of passages to measure comprehension (Johns, 1983; Beldin, 1970). Gray's assessment added to the development of the IRI as it utilized analysis of errors made while reading, which is similar to the miscue analysis in many current IRIs (Johns, 1983). Another researcher, (Wheat, 1923, as cited in Beldin, 1970) published a book in which he suggested examining reading rate and comprehension to assess reading. Beldin (1970) noted that during the 1920's, there was more

focus on the biological aspects of reading, such as the interaction between eye movements and reading, rather than on psychological factors (Beldin, 1970).

Johns (1983) notes that although the works of Waldo and Gray fueled discussion about the use of IRIs in the 1920's and 1930's, there was little written about IRIs. Instead, during this time, there was great discussion about what should be included in IRIs and what guidelines should be used for interpretations of the results of IRIs. During the 1940's, Johns noted there was increased popularity of the use of IRIs. During this time, Betts (1941, as cited in Johns, 1983) began to use an assessment he termed a "subjective reading inventory" in his reading clinic. This inventory was the first to establish some criteria for the determination of reading levels, and these levels had some limited empirical support (Killgallon, 1942, as cited in Beldin, 1970).

The 1950's saw an increase in focus on the determination of criteria for establishing independent, instructional, and frustration levels (Johns, 1983). In 1954, Betts published a book on the creation of informal reading inventories which would become a seminal work in the field of reading inventory. This book is often credited as the creation of the first formal IRI (Beldin, 1970; Johns, 1983; Fuchs, Fuchs, & Deno, 1982; Brecht, 1977; Lovitt & Hansen, 1976; Pikulski, 1990). In his book, Betts (1954) provides a detailed explanation of the purposes of creating an IRI and the advantages and disadvantages of using IRIs. In addition, Betts (1954) provides a template for creating an IRI that teachers may use when creating their own IRIs. The criteria Betts (1954) set forth for determining independent, instructional and frustration reading levels were a major contribution to the development of IRIs and continue to be used (Ekwall, 1976, Johns, 1983;).

During the 1960's, there was an increase in the publication of literature comparing IRIs to standardized achievement tests (Johns, 1983). In addition, the first studies examining the

ability of teachers to administer IRIs were published at this time (Johns, 1983). In 1965, Johnson and Kress (1965, as cited in Johnson, Kress & Pikulski, 1987) published a document which provided teachers with the background and tools needed to create and administer IRIs within their classrooms. This publication would become a resource commonly used by teachers and other researchers when developing IRIs (Anderson, 1977). The 1970s saw and expanded interest in other areas of the validity and reliability of IRIs beyond those related to the establishment of criteria for determining the differing reading levels (Johns, 1983).

Since the 1980s, literature pertaining to IRIs has primarily focused on improving the psychometric properties of various inventories (Applegate, Quinn & Applegate, 2002; Duffelmeyer, Robinson & Squier, 1989; Fuchs, Fuchs & Deno, 1982; Helgren-Lempesis & Mangrum, 1986; Nilsson, 2008; Pikulski & Shanahan, 1982, Pikulski, 1990; Rasinski, 1999; Spector, 2005; Tindal & Marston, 1996) or on creating reviews of IRIs designed for practitioners (Paris, 2002; Paris & Carpenter, 2003; Walpole & McKenna, 2006). In 1987, an update of Kress and Pikulski's 1965 document, which was designed to assist classroom teachers with creating their own inventories, was published and is still available to teachers today (Johnson, Kress & Pikulski, 1987). Commercial IRIs also have been updated in recent years to ensure they reflect current reading materials (Burns & Roe, 2002).

Psychometric properties. As informal measures, IRIs are not typically been constructed with the same focus on psychometric properties as standardized measures of achievement. Despite their informal nature, any assessment of an individual's academic skills should retain some minimal level of validity and reliability so that useful and accurate interpretation of results can be made. Therefore, a careful examination of the psychometric properties of IRIs, while mitigated by the understanding that they are informal measures, is warranted.

Validity. The source of passages and word lists used in the construction of assessments is an important consideration regarding the validity of the results they produce. Material used in the construction of IRIs constructed by teachers is typically obtained from the curriculum used by that school. Materials used to construct commercial IRIs is typically drawn from a well-established, widely available basal series (Brecht, 1977; Burns & Roe, 2002). How specific passages and excerpts from these reading series are selected to be included in the IRI varies, though most use a readability formula to determine the difficulty of the passage (Burns & Roe, 2002; Fuchs, Fuchs, & Deno, 1982; Gerke, 1980). The specific method or formulas may differ for many IRIs. As a result, there may be a difference in the reading level of passages found in different IRIs, and consequently, a comparison of the results obtained using different IRIs may be difficult (Fuchs, Fuchs, & Deno, 1982).

Thus, grade level determination made by IRIs should reflect the readability of books written at the same grade level. Some researchers have questioned this assumption, noting that the readability of text within grade levels can often vary (Bradley & Ames, 1976; Gerke, 1980). Although results of an IRI may indicate a student is capable of reading at a particular grade level, the variability found within grade level books may make generalization to other materials difficult (Bradley & Ames, 1979). Since materials for IRIs are often taken from basal readers, it is possible the variations found in the readability of books also could be found in IRIs. In this case, if the text passages at grade levels between the two IRIs are not equivalent, then comparing the grade level estimates obtained by the two IRIs will not be comparable.

Similarly, passages used within a single set of forms for an IRI also should increase in difficulty at the same rate between grade levels (Fuchs, Fuchs & Deno, 1982; Klesius & Homan, 1985; Paris, 2002). If this is not the case among passages used in IRIs, accurate judgments about

reading levels are difficult to make, and this can make comparing success or failure at different reading levels difficult (Fuchs, Fuchs & Deno, 1982; Paris, 2002).

Another concern raised about the content of passages used in IRIs includes the variable use of expository and narrative texts. This concern is based on the assumption that the way different types of texts are written may influence the rate of comprehension (Nilsson, 2008). The authors of some IRIs have deliberately included both types of texts, and some may separate types of texts between forms for comparison of comprehension based on text type (Burns and Roe, 2002; Nilsson, 2008; Spector, 2005).

Some IRIs are completely text-based (Burns & Roe, 2002), while others use pictures in conjunction with text (Anderson, 1977). For example, some provide pictures at only lower levels of reading, and some provide passages with and without pictures, allowing the examiner to determine which set to use (Nilsson, 2008). The use of pictures may increase the comprehension of a reader, so examiners utilizing IRIs which include pictures should be cognizant of this when making judgments of reading ability.

Most IRIs use a format which utilizes oral reading first to determine word accuracy and then silent reading in order to measure comprehension (Brecht, 1977; Burns and Roe, 2002), and some evidence exists that this format may influence the reading levels of students obtained on IRIs (Brecht, 1977). Gonzales and Elijah (1975) reported that rereading passages reduced the number of miscues, suggesting that the error rate of reading passages may be influenced by the reader's familiarity with a text. That is, IRIs may be assessing error rates for initial readings only, rather than true error rates (Gonzales & Elijah, 1975). However, it is important to note that rereading materials is not always possible and is certainly not efficient, so the reading ability

measured by IRIs may indeed reflect the functional ability of a reader within a real-world context.

The exact levels used to determine the independent, instructional, and frustration levels used in IRIs have been questioned. In the 1970s, several studies were conducted focusing on the comparison of frustration levels as determined by error rates on IRIs and the results of polygraph examination (Davis, 1975; Ekwall, Solis, & Solis, 1973; Ekwall, 1974). There appears to be a general agreement between the error rates used by IRIs to indicate frustration and polygraph indications of biological stress (Davis, 1975; Ekwall, Solis, & Solis, 1973; Ekwall, 1974). This relationship may be related to the types of errors counted (Davis, 1975). However, it is important to note that individuals vary on their tolerance of frustration, and there is some evidence that individuals may vary in frustration tolerance for reading based on IQ or their overall reading skill (Davis & Ekwall, 1976; Ekwall, Solis & Solis, 1973). There exists some evidence to indicate differing cut-off scores may be appropriate for different grade levels. Even when this has been the case, the changes in percentage have been minimal, ranging from 85%-98% (Klesius & Homan, 1985).

The use of only three reading types of reading levels has been criticized (Ekwall, 1976). For example, a student may expend so much energy pronouncing words correctly that they fail to attend to the meaning of the words. As IRIs use both these measures when determining grade level, it is possible for some readers to be placed in different grade levels based upon word reading accuracy than on comprehension questions, leading to confusion about what reading level best characterizes the individual. To address this problem, Ekwall (1976) proposes that the "Instructional" level be divided into two further classifications of "Definite" if both methods yield the same result and "Questionable" if both methods yield different results. Thus, teachers

would be able to target the “Definite” reading level for instruction. At least one researcher has explored the relationship between reading speed and reading levels on IRIs, finding a small correlation (Rasinski, 1999).

Similarly, the floors and ceilings used by IRIs have been questioned. Some have questioned the common practice of IRIs to use only one passage to establish reading levels (Fuchs, Fuchs, & Deno, 1982; Lovitt & Hansen, 1976).

The use of word lists to provide an estimate of grade level also has been examined. A level of 95% accuracy for word recognition for establishing an instructional level has been found to correlate with teachers’ placement of students and scores on subtests of standardized achievement measures (Fuchs, Fuchs, & Deno, 1982). In addition, graded word lists have been shown to be highly correlated with functional reading level, suggesting that the use of graded word lists in IRIs does have some predictive validity (Froese, 1976).

Research has shown mixed results for comparing IRIs with standardized achievement tests (McCabe, Margolis & Mackie, 1991; McCabe, Margolis & Barenbaum, 2001). In at least one study, achievement tests were found to overestimate reading ability (McCabe, Margolis & Barenbaum, 2001). The University of Georgia IRI has been positively correlated with the Gray Oral Reading Test and the Spache’s Diagnostic Reading Scales (DRS; Jerrolds, Callaway, & Gwaltney, 1971).

Comprehension questions are essential for determining the reading level of the reader; thus, the construction of these questions is crucial. Some have criticized IRIs for creating questions which require simply remembering and recalling information within the text, rather than higher-level thinking (Applegate, Quinn, & Applegate, 2002). There has been evidence that the use of different sets of questions may yield different results on IRIs (Peterson, Greenlaw, &

Tierney, 1978). Guidelines for creating comprehension questions for IRIs have been published (Valmont, 1972).

Passage independence, in particular, has been the focus of criticism since it cannot be easily determined if a student inferred the correct answer because of the passage or because of prior knowledge. Indeed, many questions included in IRIs have been found to contain numerous passage independent vocabulary questions (Duffelmeyer, Robinson & Squier, 1989; Marr & Lyon, 1981). In addition, IRIs demonstrate varying levels of passage independence (Marr & Lyon, 1981).

Reliability. To be an effective assessment of reading ability, assessments must demonstrate adequate reliability. An analysis of the manual of a handful of IRI manuals have led some to assert that commercial IRIs contain adequate reliability (Paris & Carpenter, 2003). Others remain critical of the reliability of commercial IRIs and suggest IRIs need improved reliability standards (Pikulski, 1982; Spector, 2005).

Because many IRIs are used to judge a student's reading level, IRIs may logically be used to monitor progress in reading ability. For this reason, most published IRIs include one or more alternate forms (Burns & Roe, 2002; Helgren-Lempesis, 1986). In addition, IRI manuals typically consider all their alternate forms to be equivalent, and leave it to the examiner to determine which forms to use during an evaluation (Burns & Roe, 2002). Therefore, alternate forms reliability is of prime importance for those who wish to use IRIs, whether it is to measure growth or for initial assessment using an IRI containing multiple forms.

The manuals of many published IRIs do not include an analysis of the of alternate forms reliability for their multiple forms (Burns & Roe, 2002; Helgren-Lempesis, 1986; Pikulski 1982; Spector, 2005). At least one IRI manual (Ekwall & Shanker, 1993) does report alternate forms

reliability and indicated adequate reliability. A review of three other IRI manuals indicated adequate alternate forms reliability (Spector, 2005). Independent research with IRIs has revealed varying results when assessing the reliability of IRIs. Empirical studies of IRIs have indicated moderate levels of reliability between alternate forms of some IRIs (Helgren-Lempsis, 1986).

However, the alternate forms reliability coefficients of some IRIs have been criticized. For example, some manuals have not reported any alternate forms reliability coefficients (Burns & Roe, 2002). Some IRIs have made a declarative statement about the reliability without including actual coefficients (Nisson, 2008). Some manuals have not included significant details about the studies the authors of IRIs have used to determine alternate forms reliability (Ekwall & Shanker, 1993). In such a case, more information would be required to make confident decisions about the reliability of alternate forms (Nisson, 2008).

In general, IRIs show great variability in the amount evidence they present for alternate form reliability. An examiner who wishes to utilize alternate forms who must choose a set of forms from multiple IRIs should carefully consider what evidence is present. Some suggestions to overcome weak alternate forms reliability when seeking to obtain a more accurate comparison of results to measure growth include using the same texts when re-testing, choosing specific texts which have demonstrated consistent grade levels, or using categorical reporting variables which are more easily compared (Paris, 2002).

Many IRIs do not report Interrater reliability (Spector, 2005; Burns & Roe, 2002). For those manuals that do report interrater reliability, high scoring consistency has been found (Spector, 2005). Empirical evidence related to interrater reliability indicates high interrater reliability can be obtained when using IRIs (Klesius & Homan, 1985). In one case, four examiners who were trained in the process of one study were found to have perfect interrater

reliability (Helgren-Lempsis & Mangrum, 1986). Although such a high interrater reliability is unlikely to be found routinely, this finding provides further evidence that high levels of interrater reliability can indeed be obtained between examiners using IRIs. Teachers being trained to administer IRIs have demonstrated similar levels of interrater reliability, even when different training methods are employed (Henney & Boysen, 1972).

However two studies utilizing multiple raters who scored a tape recorded performance of student participating in an IRI assessment found significant variability between scores (Allington, 1979; Page & Carlson, 1975). It should be noted that these studies included teachers with a wide variety of backgrounds, which may have resulted in the variability in the performance of the teachers. In addition, the Page and Carlson (1975) study used highly experienced teachers who were told simply to mark the protocols as they would if administering an IRI, rather than being given specific training on one IRI. It is likely that individuals who have been more recently received intensive, specific training in the scoring of an IRI would have higher interrater reliability scores (Pikulski & Shanahan, 1982). In general, with structured, recent training a high rate interrater reliability appears to be achievable when using an IRI.

Internal consistency analysis for comprehension questions included in IRIs has been examined. The authors of one IRI (Leslie & Caldwell, 2001 as cited in Spector, 2005) reported large standard error of measurements (SEMs) for each passage in the IRI. This indicates the student's estimated true score could vary widely for any given passage. The authors suggest that to improve the reliability of the passages, examiners should administer at least two passages per grade level (Leslie & Caldwell, 2001 as cited in Spector, 2005).

In summary, IRIs are informal measures of reading ability designed to provide examiners with an estimate of what grade level a child is reading at. As informal measures, the

psychometric properties of IRIs are often not as strong as those of standardized measures of achievement. Many IRIs report evidence of validity and generally this evidence has indicated moderate validity (Burns & Roe, 2002; Paris, 2002). Evidence of the reliability of IRIs is not as plentiful, which has led some to criticize IRIs (Pikulski, 1982; Spector, 2005). The authors of IRIs have likely not chosen to investigate the reliability of their IRIs due to the informal nature of IRIs. However, some IRIs have reported adequate reliability (Ekwall & Shanker, 1993; Nisson, 2008), suggesting IRIs have the potential to be used reliably.

Summer camps for children with autism

Three examples of summer programs or camps designed specifically for autism have been found in the literature (Hung & Thelander, 1978; Lopata et al. 2006; Lopata et al. 2008; Van Wert & Reitz, 1978). Two of the programs (Hung & Thelander, 1978; Van Wert & Reitz, 1978) focused primarily on increasing adaptive functioning behaviors and reducing problem behaviors. The remaining program (Lopata et al. 2006; Lopata et al. 2008) focused on social skills development. The programs also varied on the diagnoses of the participants and in methodology. Only one program (Hung & Thelander, 1978) was held within a traditional summer camp setting.

Over the four years of the Lopata et al. (2006) and the Lopata (2008) program was conducted it maintained the same basic format, running for six hours a day, five days a week for a total of six weeks each summer. The program curriculum was largely taken from the commercially available *Skillstreaming* (Goldstein et al. 1997) program. This curriculum was designed to increase a variety of prosocial skills through teaching, modeling, role-playing performance feedback and transfer of learning. Because the program was designed for use with multiple populations, the researchers selected only aspects of the program pertaining to the

deficits found in Asperger's Disorder to use within their program (Lopata et al. 2006; Lopata et al. 2008).

However, the program did expand the populations served and assessments used to evaluate the effectiveness of the program between the first two years and last two years of the program. The first two summers of the program included a total of 21 boys between 6- and 13-years of age with a diagnosis of Asperger's Disorder. The researchers also augmented the core social skills curriculum program with their own additional elements designed to teach the participants how to interpret idioms, non-literal statements and to understand facial expressions; though instructional these areas made up a minor portion of the total instructional time. In another addition to the curriculum, participants were randomly assigned to two conditions: one group received a behavioral treatment condition and was given tokens which could be exchanged for reinforcers when participants used skills introduced in the program and lost tokens when they engaged in problem behaviors (Lopata et al. 2006; Lopata et al. 2008).

The first two years of the program used three scales (Social Skills, Adaptability, and Atypicality) on the Behavior Assessment System for Children (BASC; Reynolds & Kamphaus, 1992) parent and teacher forms to assess for behavioral change in participants. The parents completed the scales prior to the start of the program, while staff members completed their first ratings on the 8th day of the program, and all raters completed the same ratings after the program was completed. Results indicated there was no difference in improvement between the two groups, suggesting the addition of token reinforcement to this program was not effective. Results of parent reports indicated the participants made significant improvements in social skills and adaptability as well as decreased their atypical behaviors. Staff reports were more variable, indicating a significant improvement in social skills, but no change in adaptability and an

increase in atypical behaviors (Lopata et al. 2006). The difference in staff and parent ratings is of concern when interpreting the effectiveness of the program. It is possible the staff members' initial ratings of behaviors were less accurate than the parents' ratings due to having known the participants for only eight days at the time of the first ratings. In addition, results were not contrasted with a control group and there was no evidence provided for generalization of learned social skills to other settings. Overall, these results offered positive, although tentative, support for the program.

During the last two years of this program, Lopata et al. (2008) included 54 children with high functioning autism, Asperger's Disorder, and PDD-NOS. These last two years of the program also used the same class format and the same curriculum based on the *SkillStreaming* (Goldstein et al. 1997) program and the same researcher-created instruction on non-literal language interpretation and face and emotion recognition. Also consistent with the first two years of the program, the children were divided into two groups with one group receiving points towards a reinforcer for engaging in specific pro-social skills and were penalized points if they engaged in problem behaviors while children in the other group received positive feedback for any pro-social behaviors but did not have a point system (Lopata et al. 2008).

The last two years of the program (Lopata et al. 2008) also expanded the number of outcome measures used. The program used the WISC-IV short form, three scales (Social Skills, Withdrawal, and Atypicality) and two composites (Behavior Symptoms Index [BSI] and Adaptive Skills) on the BASC, an adaptive version of the Skillstreaming Survey, the Diagnostic Analysis of Nonverbal Accuracy2 (DANVA2) and a parent satisfaction survey. As with the first two years of the program, initial assessments were completed by parents prior to the study, by

staff members on the 8th day of the program, and again by all raters at the end of the program (Lopata et al. 2008).

Parent BASC ratings indicated a significant increase in social skills and a significant decrease of withdrawing behaviors, but did not indicate a reduction in atypical behaviors, for both groups. Staff BASC ratings indicated significant improvement in social skills and a reduction in withdrawing behaviors, while suggesting an increase in atypical behaviors in the general positive feedback group and a slight decrease in atypical behaviors in the point system group. Composite scores on the Parent BASC ratings indicated a significant increase in adaptive behaviors and an overall reduction in problem behaviors for both groups. Composite ratings on the BASC completed by staff members indicated a significant increase in adaptive behaviors, but an increase in problem behaviors for the general feedback group and a slight decrease in problem behaviors in the point system group.

The Skillstreaming survey indicated significant increases in social skills for both groups. The DANVA2 was used only in the last year of the program, resulting in fewer ($n = 36$) scores. Results of the DANVA2 indicated no increased ability to interpret emotions based on facial expressions for either group. The parent satisfaction survey was only used in the first year ($n = 34$), but indicated all parents had either “positive” or “very positive” ratings for both groups.

Outcome measures from all four summers indicated increases in social skills among children with autism are attainable over the course of a summer program targeting these skills. In addition, some evidence from this program suggests reductions can be made in problem behaviors as well. Results further indicated these effects could be attained, with the exception of reducing atypical behaviors, whether specific behaviors were targeted using a token system or a more general behavioral feedback was used (Lopata et al. 2008). One limitation of this program

is the lack of comparison between all participants and a control group not enrolled in a social skills training program. For example, it is possible an increase in social skills was found simply because the participants had several opportunities to utilize social skills and to observe and interact with typically developing individuals (staff) during their day, rather than due to the social skills program. In addition, parents often desire to see positive changes in their children's behavior, and such desires may have influenced the parent ratings in this study submitted after the program.

Another example of a multi-year summer program, on a school campus, was conducted by Van Wert and Reitz (1978). This program was conducted over the course of three years on the campus of a private secondary school. The program was sponsored by a variety of sources and served 15-30 children from private and public schools each year. A staff ratio of two staff members to four or five students was maintained throughout the program. Children were divided into classrooms and each classroom was allowed to create its own schedule, but goals for every children in the program focused primarily on increasing language development, self-help skills, motor skills, and community awareness. In addition, two days a week were spent off-campus, one day spent swimming and another day spent in another community awareness activity, such as visiting a farm or a restaurant (Van Wert & Reitz, 1978).

Assessments of the children's skills were completed by referring agencies prior to each child's arrival at the program and immediately after completion of the program. After the program, feedback about successful techniques was shared with each child's parents or caregivers to encourage the continued implementation of these techniques beyond the program. The effectiveness of this program was judged using modified checklist form of the Behavior Characteristics Progression (1973) assessment. This assessment measured skill development in

24 behavioral areas, including approximately 50 sub-skills in each area. In addition, staff defined several inappropriate behaviors to target for each child and collected frequency counts on these behaviors. The authors present one child's data as evidence of the program's effectiveness (Van Wert & Reitz, 1978).

Several limitations of the Van Wert and Reitz (1978) program exist. Specifically, the authors do not present data for all the children, making conclusions about the effectiveness of the program difficult. In addition, the authors fail to provide information about total duration of the camp each summer or the total number of participants, including the number of participants attending for multiple summers, further complicating the determination of effectiveness. Due to these factors, this program provides tentative evidence, at best, of effectiveness.

Hung and Thelander (1978) conducted a 3-week residential summer camp program for 18 children campers with autism in a traditional (e.g., wilderness) summer camp setting. Campers were divided into four cabins/groups based on functional level. This program focused on four behavioral areas: self-help skills, language training, generalization of language to multiple settings and reduction of undesirable behaviors. There was a 1:1 ratio between staff and campers, an overall coordinator for each cabin, two special education teachers, two swimming instructors, a speech therapist and occupational therapist at the camp. Each staff member attended 15.5 hours of pre-camp training over the course of seven training sessions. The training focused on the principles of behavior management, shaping, and token economy (Hung & Thelander, 1978).

Individualized programs for each of the children were created soon after they arrived at camp. Staff members were required to coordinate with and seek approval from camp leaders before changing any individual behavior plans in order to maintain consistency of implementation. Staff members used physical prompts and restraints when appropriate while

implementing the behavioral plans, and campers earned tangible rewards for compliance to the behavioral programs. A summary of the individualized training programs was written up at the end of camp for the parents of each child so the programs could be extended to the home environment. The program included four classes: language, special education, swimming, and social interaction and traditional camp activities (e.g., hikes) were included at other times. The first two classes occurred in the morning, and two after lunch (Hung & Thelander, 1978).

Outcome data for this program did not include standardized measures, but instead included individual data collection for each child throughout camp. Data was not collected on the first day of the camp to allow the campers time to adjust to the new environment and routine. Baseline data was collected the next four days and treatment data was collected over the next two weeks. Behavioral reinforcers (tokens or edibles) were used for each camper to reinforce behavioral change. Overall results for the group were presented, and these indicated each child showed improvement by 15% or greater in at least one area (Hung & Thelander, 1978). These results suggest a brief, intensive summer camp program for children with autism may be successful at improving adaptive behaviors and reducing behavior problems.

Despite the positive effects of the Hung and Thelander (1978) camp, some limitations of the program exist. For example, only results for the entire group were presented, making the assessment of improvement for each of the children difficult to assess. In addition, the short duration of the camp may have influenced the production of positive results, as it is possible the behavioral plans demonstrated immediate, but not long lasting, positive behavioral outcomes. In addition, the camp appeared to offer some academic instruction (a class identified as “special education”), did not include an academic assessment, so it was not possible to determine the effectiveness of such a program on the academic functioning of the campers.

Hung (1977) conducted a three-week summer camp for children with autism that focused on question asking. This camp occurred in a traditional summer camp and included four children between the ages of 8-years, 1-month and 11-years, 9-months. The campers completed intensive, structured training in question asking within a classroom that involved training the campers using a token reinforcer to ask questions when presented with pictures or objects. The campers were then taken to other areas of the camp and onto paths through the woods where they were further reinforced for asking questions. Data from the camp suggested children were able to learn to ask questions in the classroom in response to token reinforcement, but struggled to generalize asking questions to other settings. In addition, as soon as token reinforcement was removed, all question asking also stopped (Hung, 1977). This suggests children with autism may be able to learn question asking much like many other behaviors, but learning this behavior may not generalize to other settings. This also suggests children with autism may learn a behavior but this change in behavior may not indicate a true change in functioning. That is, the reason for asking a question is to gain information, which most people find reinforcing, but if a child with autism asks the question only when he or she can gain a token, this suggests the child still does not find question asking reinforcing and thus has not had true change in functioning.

Summary

Summer camps and programs have long been considered a means to provide services to children during summer vacations when public school services are not available. Many of these programs have targeted children from low SES backgrounds or children who are underachieving in school for a variety of reasons, and these programs have shown success at increasing the academic achievement of participants (Cooper et al. 2000). Children with autism, in general, do not show a consistent profile of academic skills or weaknesses, though some common traits have

been reported. For example, children with autism tend to perform better on reading tasks which involve recalling literal and factual material or reading individual words, but struggle more than their peers to comprehend reading material, particularly if it involves abstract information or making inferences from text. Children with autism also often struggle with written language. Children with autism may be more susceptible to losing skills through the summer if they are not given opportunities to practice these skills, continuing academic instruction for children with autism during the summer months would likely be beneficial.

Several summer camps and programs have shown success at integrating children with disabilities into currently existing programs for typically developing peers (Blas, 2007; Brookman, et al. 2003). In addition, camps designed specifically for children with disabilities have proven effective in delivering therapeutic services to children with disabilities (Henley, 1996;). In addition, camps specifically for children with autism have shown success at increasing adaptive behaviors and social skills, as well as reducing problem behaviors (Hung & Thelander, 1973; Lopata et al. 1996; Lopata et al. 1998; Van Wert & Reitz, 1978). Due to these successful programs, and the likelihood of a loss of academic skills among children with autism during the summer months, additional research is needed to substantiate the effectiveness of summer camp programs for children with autism on enhancing their academic skills.

CHAPTER III

Research Methodology

The chapter contains a description of the methods utilized for data collection, participants in the study, procedures used to collect data, and instruments used. Statistical methods used to analyze the data also are detailed in this section.

Participants

A total of 28 children attended camp. One child dropped out of camp after one week due to the lengthy commute from his home to the camp, and the remaining six demonstrated behavioral problems (e.g. tantrums, refusal to interact with examiner or materials) when testing was attempted that were significant enough to prevent them from completing the assessments. As a result, only data for the 21 children who were able to complete the assessments were included in this study. Demographic information was taken from questionnaires the parents of all campers completed as a part of the application to have their child included in camp. The majority of the campers were male ($N = 17$). The campers ranged in age from 6-years, 0-months to 13-years, 1-month ($M = 9$ -years, 3-months, $SD = 1$ -year, 10-months). The majority of the campers had a diagnosis of autism ($N = 15$); the remaining were diagnosed with Asperger's Disorder ($N = 4$) and PDD-NOS ($N = 2$).

Parents reported the majority of campers were Caucasian ($N = 19$) and the other two children were described by their parents as multi-racial. Parents also reported their highest level of education. The majority of the mothers had attended college for four years or more ($N = 11$), while a few had received 1 to 3 years of college ($N = 4$), a few indicated their highest level of achievement was a high school diploma or GED ($N = 3$), and one had not obtained a high school diploma or GED. Two mothers did not provide information about their highest level of

achievement. For the fathers, the most frequently reported highest level of achievement was 4 or more years of college ($N = 8$) followed closely by those who had attended college for 1 to 3 years of college ($N = 7$), and five who reported a high school diploma or GED was their highest level of achievement. Data on the highest educational level of one father was not available. Eleven of the children resided in the same town as the camp, two lived in a large metropolitan area (population = 800,000) approximately one hour away from the camp, and one lived in a medium-sized metropolitan area (population = 250,000) approximately an hour and a half from the camp. The remaining seven children lived in communities ranging in size from very small (population = 400) to a similar size as the city in which the camp was located (population = 60,000), and all were within a one-hour drive from the camp.

Procedures

All data for this study were archival data, though the author oversaw all parts of data collection. Approval to conduct the study was granted by Institutional Review Board of Ball State University in June of 2008 (See Appendix B). Data was collected over approximately two months, beginning in early June 2008 and commencing in late July 2008.

Setting

The setting of this study (referred to as ‘autism summer camp’) was conducted at a camp facility owned and maintained by a local social service agency that provided a variety of services to children and adults with physical and mental disabilities. The camp was located on the outskirts of a Midwestern city (population = 65,000) and has served campers with disabilities for many years prior to this study. The autism summer camp was held in conjunction with weeklong residential camps that continued to be held by the organization that owned and maintained the camp, though campers in the autism summer camp rarely interacted directly with the campers

enrolled in programs conducted by the social services agency. The camp setting was a traditional setting, (e.g., wilderness) and included permanent cabins, offices, a nurses' cabin, an outdoor pool, barn, and a mess hall. In addition, the camp had an area set within the woods that included canvas tents stretched on wooden platforms, a playground, several open fields, and fire pit.

Participants for this study were sought from a support group located in the same town as the camp that targeted families who had children with autism spectrum disorders. In addition, participants were sought by advertising the summer camp to special education directors for school districts located in nearby communities. Parents who expressed interest in having their child or children attend camp then contacted the director of the camp. Interested parents were then mailed an application packet for the camp that included a cover letter explaining the goals of camp and the application procedures, an informed consent form, and questionnaires created by the director of camp requesting information regarding diagnosis, level of adaptive functioning, behavioral problems, emergency contact information, and health history. In addition, a questionnaire created by the camp director for teachers to complete regarding school behavior and academic functioning also was included. Parents were instructed that if their child was accepted to the camp they would need to provide several forms (See Appendix A), including a report of their child's most recent physical by a medical doctor, their child's most recent immunization records, most recent Individualized Education Plan (I.E.P.), and most recent psychoeducational evaluation report. These forms were used to verify each child met inclusion criteria, including verification of a diagnosis of an autism spectrum disorder made by a licensed psychologist or medical doctor. Parents were also notified that exclusionary criteria for the camp included any demonstration of self-injurious behavior.

The size of the camp facilities indicated the capacity of the camp may have required the director of the autism camp to limit the number of participants. Therefore, parents were informed through the cover letter that campers would be chosen based on a lottery system. The director of the camp reviewed all camper applications to be sure all potential participants met inclusion criteria. After careful review of all applications, two campers were determined ineligible for camp. One of the applicants was determined to be ineligible because of a lack of documentation of a diagnosis of autism, Asperger's Disorder, or PDD-NOS. The other applicant was denied admission to the camp because his parents indicated he was not toilet trained. Since some camp activities occurred in areas where a trip to a private changing facility would involve a walk of 10 minutes or more, resulting in prolonged absences from classroom instruction and participation in activities, and because of the inclusion of activities such as swimming, the camp director determined the camp setting would be inappropriate for a camper who was not toilet trained. The camp director determined the size of the camp was adequate for all the remaining applicants and therefore selection of campers did not require a lottery system.

Camp format

The campers were divided into 4 "tribes" prior to camp. The tribes were created based on age and each tribe was named after a historical Native American tribe. Each tribe had their own color and a decorated "totem pole" which served as a gathering spot for each tribe. Campers remained with their tribes throughout the day, though opportunities to interact with campers from other tribes were available throughout much of the day, with the exception of during classes, which were conducted with one tribe at a time only. Three to four counselors were assigned to each tribe. One or two additional counselors "floated" among tribes each day as needed to maintain a ratio of one counselor to each camper. In addition, a behavior specialist was available

to each tribe. Behavior specialists were all graduate-level students who were currently enrolled in special education or school psychology programs. The camp director selected each behavior specialist to ensure they had appropriate experience working with children with special needs and had completed a class in Applied Behavior Analysis.

Each camper received an individual behavior plan designed to replace a problem behavior with a more appropriate behavior or increase an appropriate behavior. For example, a behavior plan may try to replace stimulating behavior like hand flapping with taking deep breaths. Another example would be a behavior plan designed to increase eye contact. Behavior plans were designed prior to camp based on the materials provided by parents prior to camp and modified if needed, determined by the camp staff, behavioral specialists and parents, as camp continued. All counselors were expected to know the behavior plans for the campers in their tribe. In addition, at least one behavior specialist, familiar with the behavior plans for all campers, also was available for the camp each day. The specialist attended to tribes as needed to assist counselors with de-escalating behavioral outbursts and to collect behavioral data on campers. Each day, individual campers were provided with a necklace made from a string and a rectangular paper tag divided into four equal rectangles of different colors. The colors represented point levels and corresponded to a “pay-off” chart that indicated the rewards that could be earned at the end of the day. Ten points were awarded per level, and the rewards for each level increased in desirability for each level. For example, free swim in the pool was always the highest reward, and completing a puzzle was often the lowest reward, since nearly all the campers prized swimming a great deal more than completing a puzzle. Points were awarded by counselors and teachers based on specific behavioral goals listed in the children’s behavioral plans.

The basic schedule of the camp remained the same each day. Campers arrived at camp between 7:30 and 8:30 am and had free time on the camp playground until the first class. A total of four classes were held during the morning: Language Arts, Science, Math, and Art. Tribes attended classes within the same order during a given week of the camp, but the class schedule for each tribe changed each week. All classes were approximately 45 minutes in length. A 10-minute transition period was maintained between each class. In addition, a 20-minute snack break occurred after the second class. Campers ate lunch and sang group songs from 12:30 pm to 1:30 pm. During the 30 minutes after lunch, two tribes attended special programming, typically an art therapy activity provided by a local social services agency, and the campers from the other two tribes had individual tutoring. During the next 30 minutes, the tribes switched and the same programming was conducted so that each camper received the art therapy activity and individual tutoring. Next, all campers participated in a large group activity for approximately 30 minutes and prepared for their chosen pay-off activity. Pay-offs lasted approximately 50 minutes, after which campers prepared to go home for the evening.

Academic programming

Each class used disguised curriculum techniques; that is, concepts and ideas were taught using activities and games. For example, math class was held in the pool, and campers would practice counting, adding and subtracting using pool toys. Teachers created lesson plans that varied the academic skill level of the activities as necessary to ensure games and activities were appropriate for the functioning levels of the campers. For example, younger campers with less well developed achievement skills may play a counting game while other campers in the same class who had better developed skills may play a similar game involving addition skills. This ensured all campers were challenged within their academic level during classes. In addition, each

camper had a 30-minute individual tutoring session each day. Each tutor created individualized lesson plans for their campers so that their lessons were tailored specifically to the academic strengths and weaknesses of the camper.

Examiner training

Staff members of the camp were undergraduate or graduate students with training in the areas of education, special education, or school psychology. There were four teachers, one for each class subject, and all teachers were advanced education majors or local special education teachers. Teachers remained in their classroom areas during the morning when classes were conducted, but each joined an assigned tribe at lunchtime and for the afternoon activities to provide additional assistance to counselors. All staff members at the camp who would be administering assessments attended a half-day workshop in which they received specialized training in using both assessments. The author, a doctoral level student in school psychology, conducted all training and oversaw all training procedures. Training on the assessments involved a lecture on the nature of the assessments and administration procedures. Each staff member who would be administering the assessments also completed mock administrations using materials identical to those that would be used on assessments with campers. Mock administrations were conducted by each staff member with another staff members under the supervision of the author who ensured accuracy and consistency of administration for each staff member. Each staff member participated in the mock administrations as an examiner and in the role of a camper being assessed. Additional one-on-one training was provided to staff members who required remediation or requested further training. The author was available after training and during all assessments to provide further guidance and respond to questions as needed.

Assessment administration

Children with autism often demonstrate a strong preference for routines and may demonstrate behavioral problems in response to disruptions in routine (APA, 2000). In consideration of this, all assessments occurred during a 30-minute time period designated each day for individual tutoring. At the beginning of camp, this time period was selected to help children establish a routine of working one-on-one with an adult on academic tasks, as this would be the routine for individual tutoring during the remainder of camp. At the end of camp, when this routine had been established, this time period was used for assessment to prevent disruption of this routine. The number of days spent on testing depended on the progress of each child on the assessment; children who had fewer academic skills tended to ceiling out of the assessments at a faster rate than higher achieving campers, resulting in variation among campers for total test time. Test time for all campers ranged from a minimum of one day to a maximum of three days.

All participants were administered the WRAT4 (Wilkinson & Robertson, 2006) and an IRI (Burns and Roe, 2002). Both assessments used a parallel form for pre and post testing. Participants completed the Blue WRAT4 test form and Word List 1 and Passages A from the IRI during the first week of camp. Participants completed the Green Form of the WRAT4 and Word List 2 and Passages B from the IRI during the last week of camp. In the case of the WRAT4, if the assessment could not be finished in a single session, examiners were instructed to finish the subtest they were currently on and resume testing with the next immediate subtest the next day. In the case of the IRI, if assessments could not be completed in one session, examiners were instructed to finish the passage and corresponding comprehension questions and begin with the

next immediate passage at the next session. Most participants were able to complete assessments within two days, and all participants completed testing within three days.

The wilderness setting of the summer camp required examiners to complete assessments in a variety of locations, rather than a traditional setting (e.g., quiet room, at a table).

Assessments typically were conducted at picnic tables, mess hall tables, or with both camper and examiner sitting on a blanket in a field, depending on the preference of the camper and examiner. All assessments occurred a significant distance from other assessments to ensure privacy and to reduce distraction.

The author collected all materials and protocols after assessments were completed, and staff members were not longer allowed access to the assessment materials after assessments were completed. The author scored all assessments and staff members did not have access to the results of the assessments. Protocols for the WRAT4 were scored using the computerized WRAT4 Scoring Program (WRAT4-SP), whereas protocols from the IRI were hand scored according to the manual. The data gathered was entered into an SPSS data file for analysis.

Instrumentation

Wide Range Achievement Test – Fourth Edition (WRAT4)

Description. The Wide Range Achievement Test – Fourth Edition (WRAT4) (Wilkinson & Robertson, 2006) is an individually administered achievement test for individuals from 5 to 94 years of age. Administration time for the WRAT4 is approximately 15 to 45 minutes. The WRAT4 assesses basic academic skills through four subtests: Word Reading, Sentence Comprehension, Spelling, and Math Computation.

Table 3.1

Description of WRAT4 tests

Name of Test	Description
Word Reading	Reading aloud individual letters and words.
Math Computation	Completing paper-and-pencil mathematics computations.
Spelling	Spell individual words dictated by the examiner.
Sentence Comprehension	Determine appropriate word to replace missing word in a sentence.

Standard scores on the WRAT4 have a mean of 100 and a standard deviation of 10.

Standard scores are calculated for each test, and combining scores on word reading and sentence comprehension resulted in a reading composite score. The WRAT4 consists of a Blue form and Green Form, which are parallel forms, and a Combined Form. The Combined Form is equivalent to administering the Blue and Green Forms together and is used to gain a more comprehensive view of an individual's achievement abilities.

Standardization. The WRAT4 was standardized using a sample of participants that corresponded as closely as possible regarding gender, race/ethnicity, educational attainment, and geographical region to U.S. Census data collected in 2001. Standardization sites and examiners were recruited from a variety of locations within the U.S. The WRAT4 was standardized on a sample of approximately 3,000 individuals, with 100-150 individuals in 19 age categories.

Validity. The WRAT4 manual (Wilkinson, & Robertson, 2006) suggests strong validity. First, mean raw scores on the WRAT4 generally increase and decrease with age as previous research would suggest. Intercorrelations between subtests based on age were 0.60 and between grades were 0.60, indicating moderate overlap. Very minimal differential item functioning was found between WRAT4 subtest items; all correlations were 0.982 or higher. This suggests items

on the WRAT4 do not contain a bias among any of the ethnic groups included in the WRAT4 standardization sample.

The WRAT4 shows high correlation with other academic achievement measures, including the Wechsler Individual Achievement Test-Second Edition (WIAT-II), the Woodcock-Johnson III Tests of Achievement (WJ-III), the Kaufman Test of Educational Achievement: Second Edition-Comprehensive Form (KTEA-II-Comprehensive). A moderate correlation was found between the WRAT4 Word Reading subtest and the WIAT-II Word reading (0.71), the WIAT-II Decoding (0.71), WJ-III Basic Reading (0.66), and the KTEA-II Comprehensive Letter/Word Recognition (0.76). A moderate correlation also was found between the WRAT4 Sentence Comprehension and the WIAT-II Reading Comprehension (0.61), WJ-III Reading Comprehension (0.60), and the KTEA-II Comprehensive Reading Comprehension (0.42). The Reading Composite score utilized by the WRAT4 demonstrated moderate to moderately-high correlated with the WIAT-II Reading Composite (0.78), Woodcock-Johnson III Broad Reading, (0.73), KTEA-II Comprehensive Reading Composite (0.58), and KTEA-II Brief Reading (0.78).

Correlations between the WRAT4 spelling subtest with other measures of spelling were moderately-high to high: WIAT-II (0.75), WJ-III (0.89), KTEA-II Comprehensive (0.89), and KTEA-II Brief (0.78). Correlations between the WRAT4 Math Computation subtest and similar subtests on other achievement measures showed some variability: WIAT-II Number Operations (0.92), WJ-III Mathematics Calculation (0.64), KTEA-II Comprehensive Math Computations (0.75), and the KTEA-II Brief (0.73).

Divergent validity for the WRAT4 indicated the WRAT4 demonstrated higher correlations with reading subtests on other achievement tests than spelling or mathematics

subtests. Divergent validity also was found between the spelling and math subtests of the WRAT4 and the unrelated subtests of other achievement tests.

Correlations between the WRAT4 and measures of cognitive abilities also were examined. The WRAT4 demonstrated moderate correlations with Full Scale IQs obtained on the Wide Range Intelligence Test (WRIT), Wechsler Intelligence Scale for Children-Fourth Edition (WISC-IV), Stanford-Binet Intelligence Scales (SB-5), Wechsler Adult Intelligence Scale-Third Edition (WAIS-III), Wechsler Abbreviated Scale of Intelligence (WASI), Kaufman Brief Intelligence Test: Second Edition (KBIT-II), and the Reynolds Intellectual Assessment Scales (RAIS). The academic skills measured by the WRAT4 correlated higher with the verbal indices than non-verbal indices on the measures of cognitive abilities.

Effect sizes calculated between the WRAT4 for learning disabled students were moderately significant for the Reading Composite and Spelling tests and medium for the Reading, Sentence Comprehension, and Math computation subtests. Large effect sizes were found between all WRAT4 subtests for individuals with Low Cognitive abilities ($IQ < 80$) except Spelling, which had a medium effect size. Children with high cognitive abilities ($IQ > 120$) demonstrated mixed levels of correlations. Children with high cognitive abilities were higher than the control group on Word Reading, Sentence Comprehension, and the Reading composite, but did not differ from controls on Spelling and Math Computation subtests.

Reliability. Internal consistency within the WRAT4 was measured using Chronbach's alpha for the subtests and Reading Composite for all forms. Median Coefficient Alpha's for the subtests ranged from 0.87 (Math Computation, Green Form) to 0.93 (Sentence Comprehension, both forms) for all ages. Reading Composite scores demonstrated slightly higher internal

consistency, with median coefficient alphas ranging from 0.95 to 0.96. In general, internal consistency measures for the Blue and Green forms were similar.

The median Standard Error of Measurement for the Blue, Green, and Combined forms was calculated for all standard scores obtained by participants of all ages in the standardization sample. The Blue Form SEMs ranged from 4.0 (Sentence Comprehension) to 5.0 (Math Computation), and SEM of 3.0 was obtained for Reading Composite Scores. The Green Form SEMs ranged from 4.0 (Sentence Comprehension) to 5.4 (Math Computation), and an SEM of 3.0 was obtained for the Reading Composite scores. For the Combined form, SEMs were generally lower, ranging from 3.0 (Word reading and Sentence Comprehension) to 3.7 (Math Computation), and a SEM of 2.3 was obtained for the Reading Composite score.

Alternate forms reliability between the WRAT4 Green and Blue Forms was assessed using an immediate retest methodology and a delayed retest methodology. For the immediate retest condition, participants completed either the Blue or Green Form as a part of the overall standardization of the WRAT4, but then also were administered a randomly selected subtest from the alternate form of the form they were administered. Administration of the subtests was counterbalanced. Average alternate forms reliability for subtests based on age groupings was 0.86 and ranged from 0.82 (19- to 34- years of age) to 0.90 (10- to 11-years of age).

For the delayed retest condition, 330 participants from the original standardization sample were retested using a complete alternate form within approximately one month (range 8 – 86 days) after their original testing. Administration for the delayed retest condition was counterbalanced. The average alternate-form retest reliability coefficient was corrected using Guilford's (1954) formula and ranged from 0.84 (19- to 94 year olds) to 0.86 (7- to 18-year olds) and was 0.84 for all ages combined. Corrected delayed alternate-form retest reliability

coefficients for 5- to 6-years olds ranged from 0.73 (Word Reading) to 0.80 (Spelling). Practice effects, measured by calculating the mean increase in standard scores from first to second testing ranged from 0.4 to 2.2 standard score unit increases. The largest practice effects for participants 7- to 18-years of age was 2.0 standard score unit increases, which occurred on the Spelling subtest. Practice effects were greatest in the 5- to 6-year old age range, but this is likely because the items for the alternate forms in this age range are identical due to the limited achievement skills and corresponding limited means to measure the skills within this age group. Overall, results from both the immediate retest and delayed retest alternate forms reliability studies indicated satisfactory reliability for use of the Blue and Green Forms as alternate forms.

Informal Reading Inventory (IRI) (Burns & Roe, 2002)

Description. The IRI (Burns & Roe, 2002) contains two parts: Word Lists and Passages. Word Lists are designed to provide an estimate of a reader's present grade level for Independent, Instructional and Frustration reading, as well as determine a student's grade level for Listening Comprehension skills. Once an estimate of the reader's independent reading grade level has been obtained, the reader begins reading passages corresponding to that level. After each passage, comprehension questions (8 for preprimer and primer level, 10 for all higher levels) are posed to the reader. Answers to comprehension questions are deemed either correct or incorrect; samples of correct responses are included on the test protocols to compare with a reader's answer and aid in scoring. Criteria set forth in the manual indicate criteria for determining grade levels based on number of correctly read words on the Word List and the number of correctly answered comprehension questions. Administration time is usually between 20 minutes and 50 minutes.

Test Construction. The authors do not present empirical evidence regarding the psychometric properties of the IRI. The authors present information regarding the test

construction, which suggests attention to concurrent validity. The authors do not present evidence regarding reliability.

The word lists for the IRI were created using randomly selected words from vocabulary lists contained in two common basal reading series. After the word lists were compiled, participants read three levels of word lists – one designed to be at their level, one for the grade level below, and one for the grade level above. Based on these results, a word was assigned to a level if 80 percent or more of the participants from that level or higher could read a word, and less than 80 percent of the participants from the next lower level could read the word. If a word did not meet criteria using this method, it was replaced with a word that then went through the same process. After this process, completed lists were then compared to other similar, previously established lists and content was found to be highly commensurate.

Material used in the IRI was selected from the Houghton Mifflin Company's reading and literature series. The Spache Readability Formula was used to determine readability for preprimer through third grade reading passages. Fry Readability graphs were used for fourth through twelfth grade. In order to be included in the IRI, all passages were required to fit the level exactly according to the formula. Passages ranged in length between 60 and 220 words and included both fiction and nonfiction selections. Some passages were excerpts of larger passages. When an excerpt of a longer passage was used, only excerpts that made sense in the absence of the rest of the passage were utilized.

Comprehension questions were included at the end of each passage. Questions were created using guidelines set forth by Valmont (1972). Following these guidelines, the IRI comprehension questions fall into six categories: main idea, detail, sequence, cause and effect, inference, and vocabulary. Passages for preprimer through second grade levels include eight

questions, and 10 are included for passages on third through twelfth grade levels. Questions that could be answered independent of reading the passages were specifically avoided when developing comprehension questions for the IRI. However, the authors note that some vocabulary questions, by their nature, are passage independent.

During the first construction of the IRI, graduate students in a reading education program who had experience in constructing IRIs evaluated all comprehension questions. After revision, the IRI was field tested with students in the grade levels included in the IRI. Field-testing revealed the passages increased in difficulty as expected based on increasing grades and only a few questions were revealed to have problems. Reading levels as determined by the IRI were generally commensurate with the judgments of reading level made by the participants' teachers. Subsequent revisions, including the 2002 edition, went through a similar process of field-testing and revisions.

Descriptive Statistics

Descriptive statistics will be computed using SPSS 15.0 (SPSS, 2006). Descriptive statistics will be calculated for range, mean, and standard deviation of the standard scores for each WRAT4 subtest score obtained for each participant on each testing occasion. Intercorrelations for all subtests standard scores also will be calculated for each participant on each occasion. For the IRI, the range, mean, and standard deviation will be collected for the standard score on the Word List raw score and for the percentage of Comprehension Questions answered correctly.

Research Design

This study employed a pre-post test design. The design was a quasi-experimental design. All participants received the same treatment condition and no control group was used. Each

participant's scores on the WRAT4 and the IRI at the beginning of the camp will be compared with his or her scores at the end of camp. Examiners were aware of the participants' diagnoses and previous academic ability as noted on the teacher questionnaire and parent reports completed prior to camp.

Data Analysis

Data from the WRAT4 and the IRI were analyzed separately using a repeated measures MANOVA procedure. Although it is theoretically possible to use one repeated measures MANOVA to analyze results from both tests together, the small sample size of the current study is likely to reduce power significantly if such an analysis would be conducted. Therefore, separate analyses were conducted for each assessment. The analysis of multivariate data using a MANOVA procedure allowed for the comparison of multiple dependent variables and examined relationships among dependent variables. This leads to greater power to detect an effect and reduces the risk of making a type I error. However, it is important to note that this conclusion may be influenced by positive correlations between dependent variables and the effect size (Field, 2005). Tabachnick and Fidell (2007) conclude that MANOVA analysis conducted on dependent variables that are very highly and positively related is wasteful.

Variables included in the MANOVA analysis of the WRAT4 will include standard scores for all tests, excluding the Reading Composite score. The Reading Composite score is comprised of the two subsets already included in the analysis, and therefore the inclusion of this score in the MANOVA is redundant. Variables included in the IRI analysis were the number of words read correctly and percentage of comprehension questions, per type of question, answered correctly. The total number of comprehension questions posed by participants is dependent on

the number of passages read, and therefore calculating the percentage of questions answered correctly is a standardized way of comparing participants' scores.

Sample size

Stevens (1980) indicates the power of a MANOVA is small for analyses with small-to-moderate group and effect sizes. Tabachnick and Fidell (2007) indicated it is necessary to have more cases than dependent variables in every cell to have adequate power. Thus, when conducting MANOVA analysis on data collected in this study, careful examination of the sample size as well as effect size will need to be undertaken to ensure the analysis is appropriate for the data.

Fit Statistics

Levene's test was utilized to assess the assumption of equal variances. Post hoc tests were conducted for all significant results. Effect sizes for all significant results also were conducted. In addition, Roy-Bargmann step down procedures were used to determine the contributions of the various dependent variables to significant effects (Tabachnick & Fidell, 2007).

CHAPTER IV

Results

This chapter reports the results of the current study using the Wide Range Achievement Tests – Fourth Edition (WRAT4; Wilkinson & Robertson, 2006) and the Informal Reading Inventory (IRI; Burns & Roe, 2002). Results for each measure are reported separately. Descriptive statistics for both measures are reported. MANOVA results and sign test results for the WRAT4 are reported. Qualitative interpretations of the IRI are reported.

Research Hypothesis

The first research question considered the effectiveness of the academic remediation of the eight-week academic and behavioral remediation summer camp as measured by the WRAT4:

Research Hypothesis: Children with autism who participate in a summer camp including academic instruction will maintain or improve their reading, spelling, and mathematical computation skills as measured by the Wide Range Achievement Test – Fourth Edition.

Descriptive Statistics

The means, standard deviations, ranges and 95% confidence intervals for the standard scores on the subtests on the WRAT4 were calculated for each participant. Scores are based on a standardized mean of 100 and a standard deviation of 15. Scores between 85 and 115 are considered to be in the Average range. The first administration of the WRAT4 was the Blue Form, and the second administration was the Green Form.

The means, standard deviations, and 95% confidence intervals for the four subtests of the WRAT4 are presented in Table 4.1. The means for both the first and second administration of the WRAT4 were within the Average range on the Word Reading test and were within the Below

Average range on the Math Computation, Spelling, and Sentence Comprehension tests. It is important to note that the standard deviations for all tests were above 15 points, indicating significant variability between children in the sample. Mean scores on the WRAT4 generally showed stability between the first and second administrations for this sample.

Table 4.1

Descriptive Statistics for Standard Scores on the Wide Range Achievement Test – Fourth Edition (WRAT4).

Name of Test	Minimum	Maximum	Mean	SD	95% confidence interval
First administration					
Word Reading	57	145	95.00	25.49	82.33-107.67
Math Computation	55	110	78.28	16.90	69.87-86.68
Spelling	55	129	86.61	21.67	75.83-97.39
Sentence Comprehension	62	120	84.78	16.07	76.79-92.77
Second administration					
Word Reading	67	145	94.11	23.38	82.48-105.74
Math Computation	55	109	80.44	18.93	71.03-89.86
Spelling	23	114	85.06	24.05	73.10-97.01
Sentence Comprehension	67	118	89.44	17.20	80.89-98.00

Note. N = 21

MANOVA results

A repeated measures MANOVA was conducted to determine if there was a significant change in WRAT scores over time, as well as whether there were differences in scores by gender, and an interaction between time and gender, which would suggest differences in change over time for males and females. The alpha level was set at 0.05. The MANOVA indicated there was no main effect for gender ($F = 0.53$; $p = 0.81$). The MANOVA also indicated there was no significant interaction between gender and time ($F = 0.63$; $p = 0.65$). Due to these findings, all subsequent results are reported without regard to gender. In terms of time, the MANOVA indicated there were no significant changes between the first administration and the second administration ($F = 0.44$; $p = 0.44$).

Sign Test

Given the complex nature of the MANOVA analysis, the power of a MANOVA run on this data was likely lowered significantly due to the small sample size (Tabachnick & Fidell, 2007). Therefore, an alternative non-parametric sign test also was conducted on each measure to investigate changes in scores between the first and second administration. The Sign Test may be more appropriate given the small sample size that characterizes the current research (Higgins, 2004). A Bonferroni correction of $0.05/4 = 0.0125$ was used to account for the inflated alpha resulting from multiple comparisons. The multiple comparisons involve comparing the different tests of the WRAT4 with each other. None of the subtests evidenced statistically significant change (Table 4.2). However, the Sentence Comprehension subtest evidenced the greatest change and was very nearly a significant change. The Cohen's d effect sizes computed for each subtest confirmed there was little effect of change over time for the Word Reading, Math

Computation, and Spelling subtests, but indicated at small effect for Sentence Comprehension based on Cohen's guidelines for interpreting d (Cohen, 1988).

Table 4.2

Sign Test Results for first and second administrations of tests on the Wide Range Achievement Test – Fourth Edition (WRAT4).

Name of Test	M	P	D
Word Reading	4	0.10	0.09
Math Computation	1	0.83	0.05
Spelling	0.5	1.00	0.03
Sentence Comprehension	5.5	0.02	0.35

Research Question

The research question considered the effectiveness of the academic remediation of the eight-week academic and behavioral remediation summer camp as measured by the IRI:

Research question: Will children with autism who participate in a summer camp including academic instruction improve the functional, instructional, frustration reading levels and the listening comprehension level as measured by an Informal Reading Inventory?

Descriptive Statistics

Results of the IRI are computed in terms of grade level for instructional, independent, and frustration reading levels, as well as for listening comprehension. Because the measure is not normally distributed, standard parametric statistics are not appropriate for this measure. In addition, the variable to be measured, grade level, is ordinal rather than continuous, also making

the use of parametric statistics inappropriate. Instead, qualitative analysis of increases or losses of reading level for each level were calculated. A total of 18 campers were able to complete the first and second administration of the IRI. Campers who were not able to complete both administrations were excluded from analysis.

Table 4.3 presents the number of students whose scores improved, declined, or remained unchanged in their word list reading abilities between the first and second administrations of the IRI. For the independent word list reading level of the IRI, 1 student improved, 3 students declined, and 14 students remained unchanged. For the instructional word list reading level of the IRI, 3 students improved, 5 students declined, and 10 students remained unchanged. For the frustration word list reading level of the IRI, 2 students improved, 5 students declined, and 11 students remained unchanged.

Table 4.3

Number of students who improved or declined as measured by word list for the IRI.

Level	Improved	Declined	Unchanged
Independent	1	3	14
Instructional	3	5	10
Frustration	2	5	11

Table 4.4 presents the number of students who improved, declined, or remained unchanged in their passage comprehension reading abilities between the first and second administrations of the IRI. For the independent passage reading level of the IRI, two students improved, three students declined, and 13 students remained unchanged. For the instructional passage comprehension reading level of the IRI, five students improved, four students declined,

and nine students remained unchanged. For the frustration passage comprehension reading level of the IRI, seven students improved, five students declined, and six students remained unchanged. For the listening comprehension passage comprehension reading level of the IRI, seven students improved, five students declined, and six students remained unchanged.

Table 4.4

Number of students who improved or declined as measured by passage comprehension for the IRI.

Level	Improved	Declined	Unchanged
Independent	2	3	13
Instructional	5	4	9
Frustration	7	5	6
Listening Comprehension	7	5	6

Qualitative Analysis

The IRI scores demonstrated remarkable variability compared to the WRAT4 scores. That is, participants tended to score similarly on the first and second administrations of the WRAT4, but many participants did not score similarly between the first and second administration on the IRI. Therefore, a qualitative examination of the scores for four children on the WRAT4 and the IRI was undertaken to explore systematic differences between the children that may provide an explanation for the variability in their IRI scores.

Children 1 and 2 were selected based on gender and because their ages were within one year of each other. Figure 4.1 presents Child 1's WRAT4 and IRI scores. Child 1 was a Multi-racial female diagnosed with autism who was 8-years, 7-months of age at time of pretest. As can

be seen from the WRAT4 scores, Child 1 demonstrated similar scores on all four WRAT4 tests. Child 1 improved slightly on Word Reading, Sentence Comprehension, and Math Comprehension, but remained consistent on the Spelling test. The IRI scores for child 1 indicate improvement in reading individual words from the Word List on the IRI for both the Independent reading level and the Frustration Reading level and remained consistent on the Instructional level. However, on passage comprehension, Child 1 declined markedly on the second administration of the IRI for all reading levels.

Figure 4.1. WRAT4 and IRI scores for Child 1.

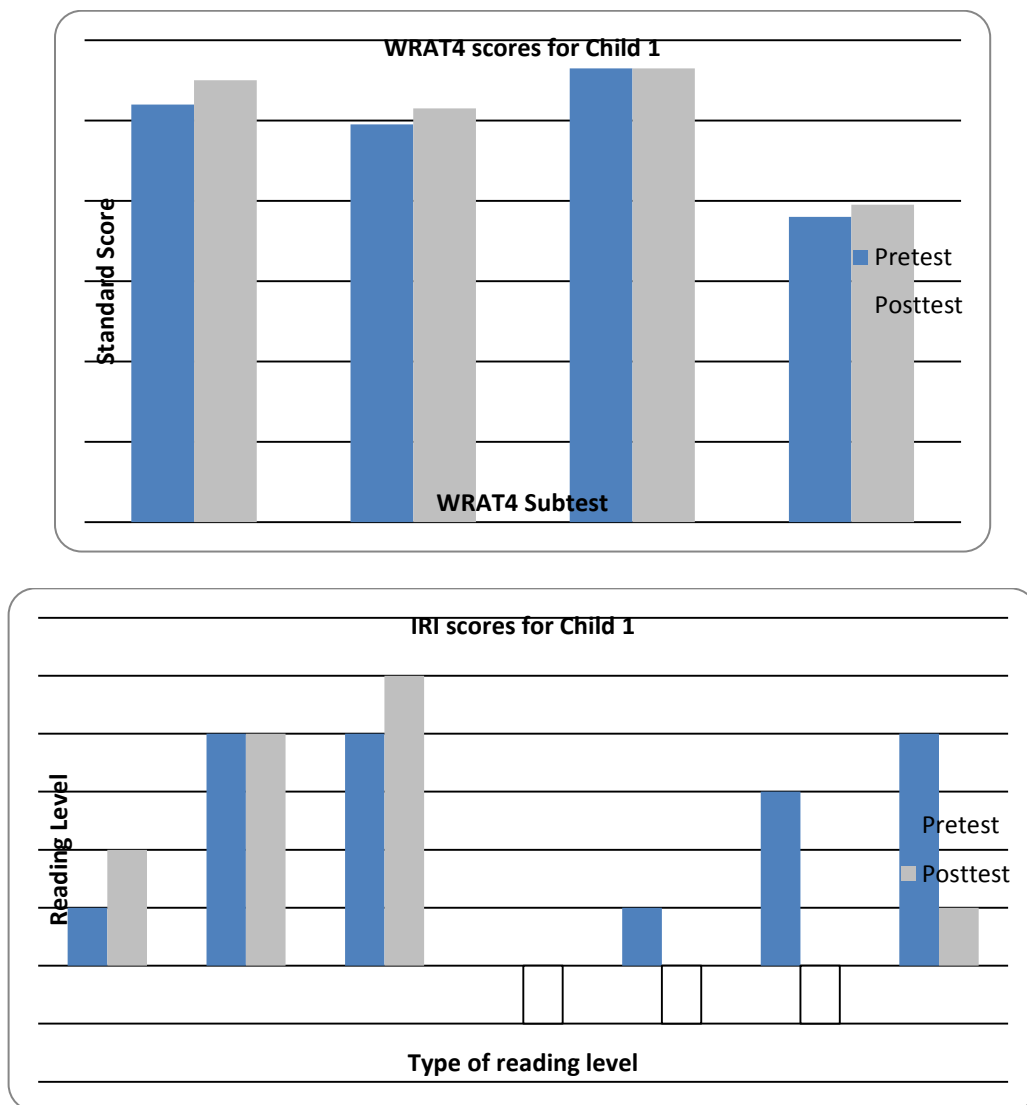
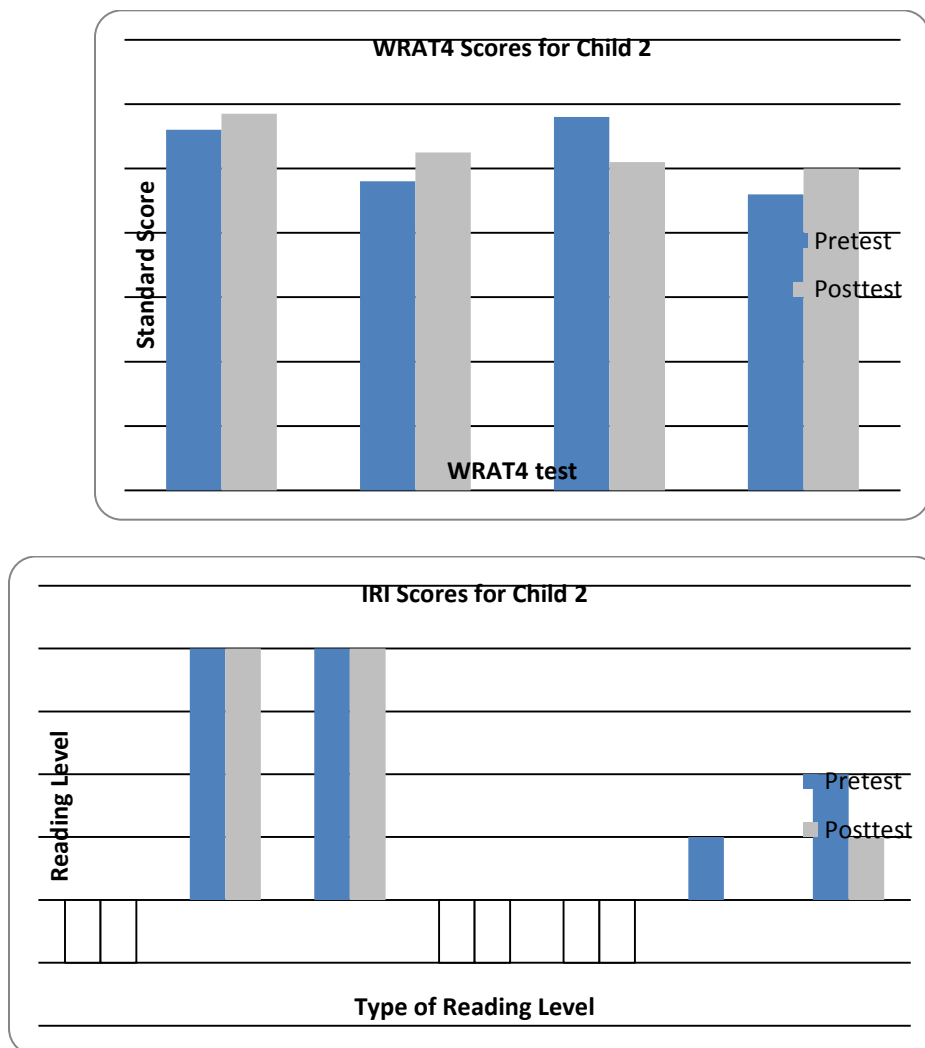


Figure 4.2 presents WRAT4 and IRI scores for Child 2. Child 2 was a Caucasian female diagnosed with autism who was 7-years, 7-months of age at time of pretest. Child 2 demonstrated slight improvement on the Word Reading, Sentence Comprehension, and Math Comprehension tests, and declined slightly on the spelling subtest. Child 2 displayed a similar performance on the Word List of the IRI. Child 2 demonstrated similar reading levels on the Independent and Instructional reading levels on the Passage Comprehension portion of the IRI, but declined on both Frustration and Listening Comprehensions on the Passage Comprehension of the IRI

Figure 4.2. WRAT4 and IRI scores for Child 2.



Children 3 and 4 were chosen based on gender and because their ages were within one year of each other. Figure 4.3 presents scores for Child 3 on the WRAT4 and IRI. Child 3 was a Caucasian male diagnosed with autism who was 9-years, 5-months old at the time of pretest. Child 3 demonstrated improvement on Word Reading, Spelling, and Math Computation, and remained consistent on Sentence Comprehension. Child 3 demonstrated a consistent performance on the Word List section of the IRI. Child 3 demonstrated a similar performance on the independent level of the Passage Comprehension of the IRI and improved on all other types of reading levels.

Figure 4.3. WRAT4 and IRI scores for Child 3

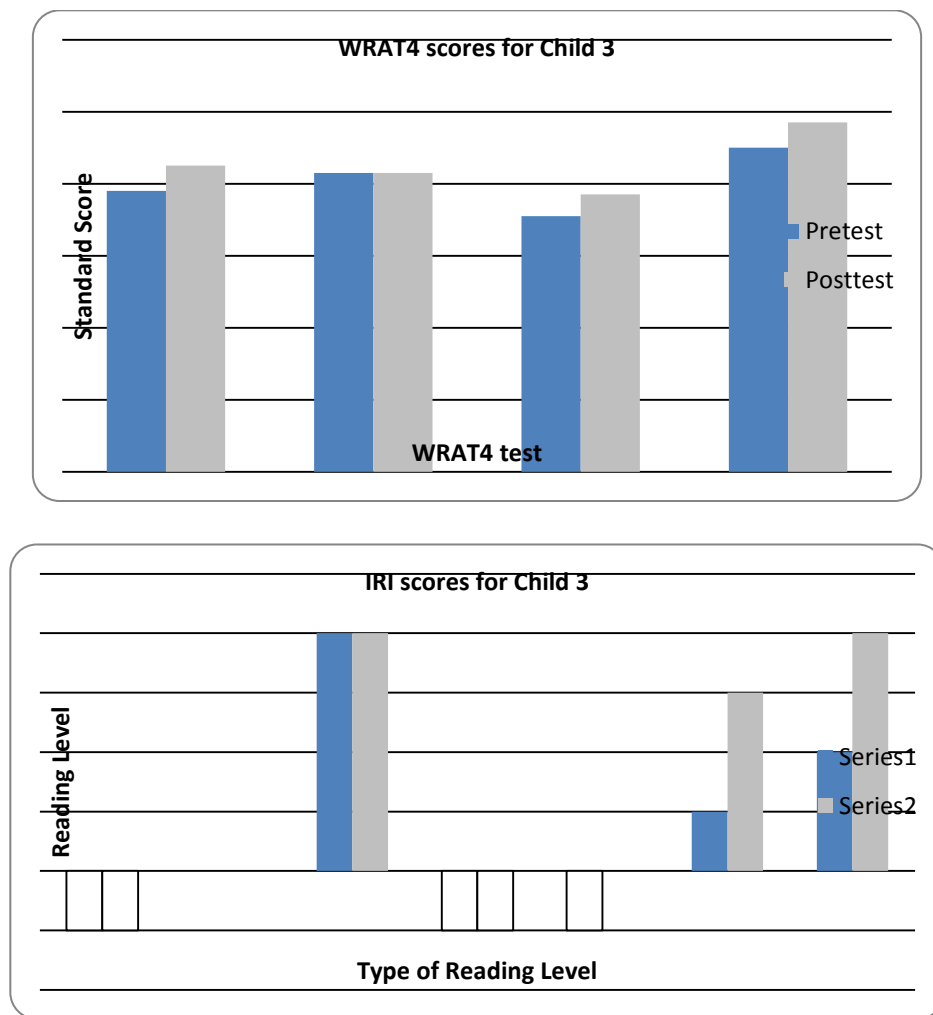
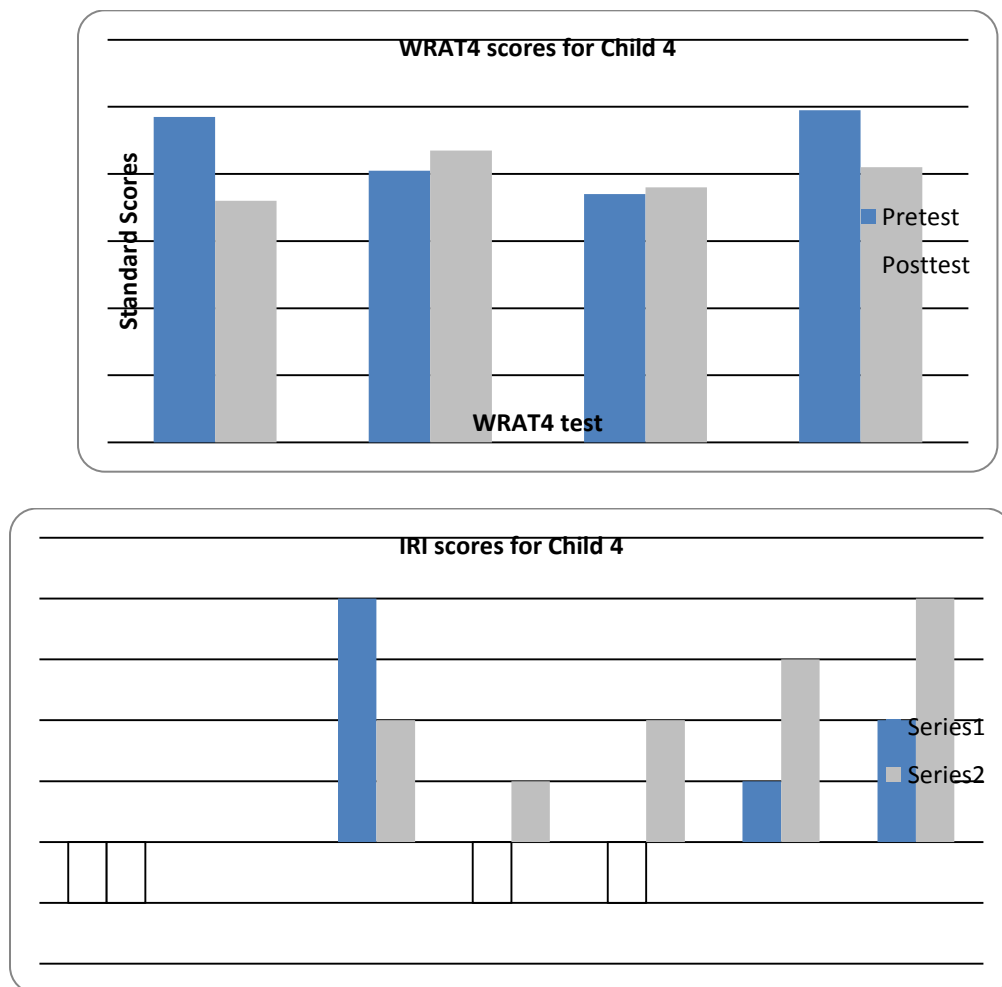


Figure 4.4 presents Child 4's scores on the WRAT4 and IRI. Child 4 was a Caucasian male diagnosed with autism who was 10-years, 4-months old at the time of pretest. On the WRAT4, Child 4 showed variability in scores between the first and second administration. Child 4 declined on Word Reading and Math Computation, but improved on Sentence Comprehension and Spelling. Child 4 demonstrated consistency on the IRI Word Lists, displaying similar performance on Independent and Instructional reading levels, but declining slightly on Frustration reading level. However, Child 4 made significant improvement on the IRI Passage Comprehension on all types of reading levels.

Figure 4.4 WRAT4 and IRI scores for Child 4



A comparison of the four children shows all four children demonstrated reasonable consistency over time on the WRAT4. However, all four children demonstrated different patterns of performance on the IRI over time. Only Child 4 demonstrated marked improvement on the second IRI; Child 1, Child 2, and Child 3 remained consistent, declined or improved in only two areas. Child 1 and 2, who were both female, demonstrated different patterns of performance over time on the IRI. Child 3 and Child 4 who were both males, also demonstrated differences in IRI performance over time. Child 4, who improved the most on the second administration of the IRI, was the oldest child of the 4, but was still within one year of Child 3, suggesting age was not a factor in his improvement.

Analysis of IRI Error Rates

The error rates for the IRI represent the percentage of each type of question on the IRI that was answered incorrectly. Answering each type of question requires different skills; for example, answering vocabulary-based questions requires rote memory while answering questions involving inference requires abstract thinking. An analysis of the abilities of children with autism to answer different types of questions may reveal unique reading skill patterns for children with autism. The means and standard deviations of error rates for each question type for oral reading and listening comprehension are reported in Table 4.5 On the first administration of the IRI, the means for error rates of all questions are similar, with the exception of Inference questions which have a much higher error rate. Thus, this type of question proved more difficult for participants to answer on the first IRI. Examination of the error rates for the second administration of the IRI indicates a significantly lower error rate for inference questions. This suggests that inference questions were easier for participants to answer on the second

administration, suggesting an improvement in the ability to answer inference questions over time.

Table 4.5

Means and standard deviations of error rates for each question type on IRI.

Question type	Oral Reading		Listening Comprehension	
	Mean	SD	Mean	SD
First Administration				
Main idea	39.22	38.70	72.22	42.78
Sequence	47.17	35.45	88.89	27.42
Detail	44.00	33.82	79.17	32.95
Cause and effect	47.47	37.57	67.61	38.12
Inference	62.83	35.41	64.00	36.83
Vocabulary	40.29	37.95	67.83	41.81
Second Administration				
Main idea	34.17	35.57	61.11	50.16
Sequence	48.61	32.89	91.67	25.73
Detail	46.22	26.66	62.78	35.55
Cause and effect	50.17	30.79	88.89	32.34
Inference	37.22	28.71	46.28	46.99
Vocabulary	30.39	40.88	60.00	50.71

Table 4.6 presents the results of paired-samples *t*-tests comparing the error rates for the different types of questions on the IRI and reports the effect sizes of those differences. The assumption of normality was not met. A Bonferonni correction was used to account for multiple comparisons, resulting in a significance level of 0.008. The Vocabulary subtest does not contain vocabulary words questions for the preprimer passages on the IRI. Therefore, students who have a frustration level at the preprimer level are not included in this analysis, and this accounts for the differences in degrees of freedom for the vocabulary test.

Table 4.6

Paired samples t-test for error rates on IRI between first and second administration.

Question type	Oral Reading				Listening Comprehension			
	<i>t</i>	df	<i>p</i>	<i>d</i>	<i>t</i>	df	<i>p</i>	<i>d</i>
Main idea	0.512	17	0.62	0.12	0.720	17	0.48	0.17
Sequence	0.153	17	0.88	0.04	0.294	17	0.77	0.07
Detail	0.271	17	0.79	0.06	1.420	17	0.17	0.33
Cause and effect	0.245	17	0.81	0.06	1.580	17	0.13	0.37
Inference	2.812	17	0.01	0.66	1.940	17	0.07	0.46
Vocabulary	0.526	13	0.61	0.14	0.710	14	0.95	0.02

The differences in error rates by question type for oral reading approached significance for questions dealing with inference ($t=2.812$, $p=0.01$). The Inference question type for listening comprehension also approached significance ($t=1.940$, $p=0.07$). Based on Cohen's guidelines for

interpreting d (Cohen, 1988), all but one of the effect sizes indicated no effect or a small effect; the effect size for inference questions for oral reading indicated a medium effect size.

CHAPTER V

Discussion

This chapter further discusses the results of the current study and its relevance to past and future research. In addition, this chapter will review the rationale of the study which was presented in Chapter I and the results of the current study be discussed in relation to the two research questions presented in Chapter I. The implications of this study to the field of assessment and intervention for children with autism also will be discussed. Finally, limitations of this study and recommendations for future research will be presented.

Rationale for the study

The purpose of this study was to examine the effectiveness of academic remediation conducted during a six-week academic and behavioral remediation summer camp for children with autism. Summer camps have been used to deliver services to children with many types of psychological and developmental disorders (Bennett, Cardone & Jarczyk, 1997; Davis-Berman & Berman, 1989; Eikeanaes, Gude & Hoffart, 2006; Forehand, Mulhern & Rickard, 1969; Gillis & Simpson, 1991; Heckel, Hursh & Hiers, 1977; Henley, 1999, Hinshaw, 2002; Hung & Thelander, 1978; Kelley, Coursey & Selby, 1997; Larson, 2007; Lattal, 1969; Romi & Kohan, 2004; Russell, 2003; Stich & Senior, 1984; Yssel, Margison, Cross & Merbler, 2005). Some children and adolescents may respond better to or prefer receiving therapeutic services in a summer camp setting rather than through traditional talk therapies (Davis-Berman & Berman, 1994; Glass & Myers, 2001; Russell, 2001). Given the communication and social deficits which are a hallmark of autism, summer camps may provide an ideal setting with which to provide therapeutic services for children with autism.

One example of such a summer camp (Hung & Thelander, 1978) conducted in a wilderness setting that sought to provide treatment for children with autism indicated significant improvements for their participants, though this study primarily examined adaptive skills. A different summer program, though not conducted in a wilderness setting, also found success in improving the adaptive skills of children with autism (Van Wert & Reitz, 1978). Another summer program not held in a wilderness setting found success improving the social skills of children with high-functioning autism and Asperger's Disorder (Lopata, Thomeer, Volker & Nida, 2006; Lopata, Thomeer, Volker, Nida & Lee, 2008). Although these studies indicate summer programs and summer camps can be successfully used to deliver therapeutic services to children with autism, none of the above studies examined core academic skills. Skills such as reading are important for independent living (e.g., reading cooking directions and safety warnings) and therefore, should be considered a component of treatment for children with autism. The primary focus of this study was to investigate the academic skill levels of children with autism, as measured by the Wide Range Achievement Test – Fourth Edition (WRAT4; Wilkinson & Robertson, 2006) and an the Informal Reading Inventory (IRI; Burns & Roe, 2002) prior to attending camp and after having received services at the camp.

Research Hypothesis

The research hypothesis examined the effectiveness of academic remediation at the summer camp using the WRAT4 (Wilkinson & Robertson, 2006). Participants were assessed using all tests on the WRAT4 prior to participating in the camp and at the completion of camp. The WRAT4 was selected primarily due to the standardized nature of the assessment. This allowed scores for the participants to be accurately compared using pre- and post-test methodology. In addition, the standardized nature allowed the participants, who were of different

ages and grade levels, to be compared to each other to measure the effectiveness of the camp overall. The WRAT4 also was selected due to its ease of administration and because it contained measures in three essential areas of academic achievement: reading, mathematics, and spelling.

The means, standard deviations, and 95% confidence intervals obtained on the WRAT4 indicated significant variability in the skill levels of the participants. Given the heterogeneity of functional levels in children with autism (APA, 2000), the variability of scores obtained on the WRAT4 is an unremarkable finding. The overall scores for participants were in the Average range for the Word Reading test and in the Below Average range on Math Computation, Spelling, and Sentence Comprehension tests.

In other studies examining achievement skills among children with autism, reading scores were not directly compared to other achievement skills; however, mean overall reading standard scores reported did not appear significantly greater than other achievement skills (Mayes & Calhoun, 2003a; Mayes & Calhoun, 2003b). Minshew et al. (1994) also reported reading scores that did not appear significantly different from other achievement skills using the K-TEA, though again they did not directly compare scores. However, some studies have suggested a difference between word identification and reading comprehension among children with autism (Church et al. 2000; Frith & Snowling, 1983; Minshew et al. 1994; Mayes & Calhoun, 2003b; Nation et al. 2006; Rumsey & Hamburger, 1990; Whitehouse & Harris, 1984).

The current study utilized the Word Reading and Sentence Comprehension tests on the WRAT4 to measure reading. The Word Reading test requires the use of decoding and identification of individual words whereas the sentence comprehension test requires the integration of word recognition skills and the use of context clues to determine words omitted from sentences. Children in the current study generally performed better on the Word Reading

test than the Sentence Comprehension test. The finding that Word Reading skills of participants were greater than their Sentence Comprehension skills is consistent with past research conducted on the reading abilities of children with autism which suggested reading individual words to be a strength compared to reading comprehension for many children with autism (Church et al. 2000; Frith & Snowling, 1983; Minshew et al. 1994; Mayes & Calhoun, 2003b; Nation et al. 2006; Rumsey & Hamburger, 1990; Whitehouse & Harris, 1984). The findings of this study further suggest children with autism present an Average skill level with reading individual words but may struggle, by comparison, with reading comprehension.

Mayes and Calhoun (2003a) found participants' spelling skills were commensurate with their overall cognitive abilities. However, the average cognitive abilities, and therefore the mean spelling scores, reported for this study varied significantly among the sample, ranging from the Very Low range to the Superior range (Mayes and Calhoun, 2003a). The present study did not measure cognitive abilities; however, a range of scores from the Low range to the Average range was found in this study, which appears consistent with the findings of Mayes and Calhoun (2003a). Minshew et al. (1994) reported spelling scores that were consistent with other achievement skills. Results from the current study indicated spelling scores on the WRAT4 did not significantly differ from other achievement skills, consistent with the findings of Mayes and Calhoun (2003a) and Minshew et al. (1994).

In the area of mathematics, some researchers (Gagnon et al. 2001) have hypothesized that the unique deficits in processing which favor literal and factual learning in autism may suggest children with autism would show superior skills compared to their typically developing peers in the areas of mathematics, though this has not been demonstrated in research (Church et al. 2000; Gagnon et al. 2001; Goldstein et al. 1994; Griswold, 2002; Minshew, 1994; Minshew 1992). In

addition, research has not suggested children with autism have greater mathematical skills than their overall cognitive abilities (Mayes & Calhoun, 2003a; Mayes & Calhoun, 2003b). The present study did not compare cognitive abilities to achievement skills, but found mathematical skills consistent with other achievement skills. A control group of typically developing peers was not included in the present study; however, scores for the participants in this study had scores in the Below Average range, suggesting slightly lower mathematical abilities than their typically developing peers.

No studies examining the effects of academic remediation for children with autism have been found in the literature; therefore, comparison of the results of the pre- and posttesting with the WRAT4 in the present study with other studies in the literature is not possible at this time.

MANOVA tests comparing WRAT4 scores indicated no significant differences based on gender and no interaction between gender and time. This suggests effects of academic remediation were not different for female or male participants. Overall, MANOVA results indicated scores on the WRAT4 did not change significantly from pre- to posttesting. This suggests the academic remediation at the camp did not significantly improve participants' academic achievement in reading, spelling, and math as measured by the WRAT4.

Although results indicated the academic remediation at the summer camp did not improve academic functioning for participants, it is important to note a decline in achievement abilities was similarly not observed. Some researchers have reported the phenomenon of "summer learning loss" in which students lose academic achievement abilities over the summer (Alexander et al. 2007; Cooper et al. 1996). Cooper et al. (1996) suggested summer learning loss can equal a full month of instruction. Others have noted there is variability in the degree of summer learning loss based on the SES of the student, indicating children with low SES lose

skills during the summer while those from higher SES backgrounds maintain or gain skills over the summer (Alexander et al. 2007). Cooper et al. (1996) and Alexander et al. (2007) did not include children with autism in their studies, making the degree of summer learning loss among children with autism difficult to predict. Despite this, the finding of no significant loss in skills suggests the academic remediation at camp may prevent summer learning loss and did not have a negative impact on the achievement skills of the participants.

Due to the complexity of the MANOVA test, analyses on small samples such as those in the current study may reduce the power of the test. Therefore, a Sign Test (nonparametric test) was conducted to look for differences in WRAT4 scores over time. Sign Test results were similar to the MANOVA results and confirmed there was no significant change over time on any of the tests included in the WRAT4. However, it is important to note change in scores on the Sentence Comprehension test approached significance for pre- to posttest. A comparison of means for pre- and posttest scores indicated participants obtained higher scores on the Sentence Comprehension test at the end of camp. Effect sizes indicated a small effect for change among the sentence comprehension scores according to Cohen's (1988) guidelines. This suggests participants slightly improved their sentence comprehension skills during camp. Given the report by several researchers that reading comprehension is a skills deficit for children with autism (Church et al. 2000; Frith & Snowling, 1983; Minshew et al. 1994; Mayes & Calhoun, 2003b; Nation et al. 2006; Rumsey & Hamburger, 1990; Whitehouse & Harris, 1984), this finding of even a slight change in this area is notable.

Research Question

The research question examined the effectiveness of academic remediation at the summer camp using the IRI (Burns & Roe, 2002). One objective for using the IRI, in addition to the

WRAT4, was because the WRAT4 measured reading comprehension using sentence comprehension, which may not accurately represent the participants' ability to read longer passages, such as those typically found in textbooks and story books. In addition, the IRI gave more information about reading levels for each participant, including measuring his or her Independent, Instructional, and Frustration reading levels.

One study was found in the literature that used the Classroom Reading Inventory (CRI; Silvaroli, 1993), an Informal Reading Inventory, with children with Asperger's Disorder (Myles et al. 2002), although it is important to note this study did not include children diagnosed with autism. The Myles et al. (2002) study compared instructional reading levels with actual grade levels of participants. Results of the Myles et al. (2002) study indicated no significant difference between participants and their grade levels, suggesting children with Asperger's Disorder do not typically present with deficits in reading comprehension.

The present study measured change in reading levels for both word lists and passage comprehension from pre- to posttest. Results from the current study indicated significant variability among IRI scores for several children from the pre- and posttest for both the word list and passage comprehension portions of the IRI, though variability on the passage comprehension portion of the IRI was significantly greater than on the word list part. The rates of improvement and of decline on the IRI appeared very similar for all reading levels. That is, there did not appear to be a trend in children improving or declining in reading level for any type of reading level or between word lists and passage comprehension. This finding suggests the results of the IRI used in this study (Burns & Roe, 2002) are not reliable, which is consistent with criticism by several writers that IRIs do not make adequate attempts to establish test-retest reliability (Pikulski, 1982; Spector, 2005).

It should be noted that the lack of reliability found in the current study could be related to the behaviors commonly found among children with autism. That is, given the behaviors often displayed by children with autism, it is possible some children exhibited inattentiveness, self-stimulating behaviors, echolalia and focal interests during one testing session but did not exhibit those behaviors during the other testing session. Such discrepant behaviors could account for gains or losses on the IRI. Although the WRAT4 showed significant stability from pre- to posttest, the IRI was administered after the WRAT4, and it is possible behavior problems such as inattentiveness became progressively worse throughout testing, having a greater impact on the scores of the IRI rather than the WRAT4.

Due to the contrast between the consistency of results obtained on the WRAT4 and inconsistent results obtained on the IRI, a qualitative analysis of four individual participants was undertaken to explore for possible explanations as to why the scores for some participants changed on the IRI while those of other participants did not. In addition, this analysis allowed for comparison of scores on the WRAT4 by a specific participant with scores obtained on the IRI.

The first two children were selected because they were both females who were within one year of age of each other. Child 1 evidenced significant decline in IRI scores at posttest, while Child 2 demonstrated consistency on all levels of the IRI with the exception of a slight decline in frustration and listening comprehension levels for passage comprehension. Both Child 1 and Child 2 evidenced slight improvement on the WRAT4 from pre- and posttest, with the exception of Child 2 declining slightly on the Spelling test. Therefore, although both children evidenced slight improvement in academic skills in most areas on the WRAT4, including reading scores, the IRI indicated very different results regarding reading. Given the similarities in age, it is unlikely that the differences in reading achievement for these two participants are related to age.

Similarly, both participants were female, so gender would not account for these differences. It is possible behavior played a role in the differences of the IRI; that is, it is possible Child 1 had significantly greater behavior problems during posttest of the IRI than Child 2. However, this is unlikely since similar gains on the WRAT4 were noted for both students. Although SES was not specifically measured for participants, a review of records indicated Child 1's mother had a High School diploma or GED and her father's educational level was not reported. Records for Child 2 indicated her mother had four or more years of college, while her father had one-three years of college. Alexander et al. (2007) reported a significant gain in achievement scores for children from higher SES household; therefore, it is possible the differences in educational levels between the parents of the two children could account for the differences in reading skills on the IRI for the two children. For example, Child 2's parents, having a higher level of educational achievement, may have provided a more enriched environment for Child 2 and/or tutored Child 2 at home, whereas Child 1's parents may not have, resulting in less decline for Child 2 with regards reading.

Child 3 and Child 4 also were selected based on gender and age; both were males with birthdates within one year of each other. Child 3 demonstrated consistency in scores on the word reading portion of the IRI and on the Independent reading level of passage comprehension and improvement on the Instructional, Frustration and Listening Comprehension levels for the passage comprehension portion of the IRI. Child 3 demonstrated similar performance on all levels of the IRI except an increase on Frustration and Listening Comprehension levels of the IRI. Child 3 slightly improved on all subtests of the WRAT4, with the exception of demonstrating a consistent performance stable on Sentence Comprehension. Child 4 improved on Sentence Comprehension and Spelling, but declined slightly on Word Reading and Math

Computation. Thus, despite a similar pattern of slight gain in achievement on the IRI, Child 3 and Child 4 demonstrated slightly different patterns of achievement on the IRI.

A review of records indicated Child 3's parents both had one-three years of college education, while Child 4's mother had four or more years of college and his father had three years of college. This finding is not commensurate with Alexander et. al (2007)'s finding that higher SES would likely lead to more gains, as based on those finding it would be expected Child 4 would improve more than Child 3. A further review of records indicated Child 3's parents listed his autism as "moderate" while Child 4's parents listed his autism as "mild." It is possible that Child 4 made greater gains in reading comprehension due to having a milder autism. However, it should be noted that the severity of autism was reported for informational purposes for the camp only; there are no established guidelines for such descriptions of autism, nor was the severity of autism confirmed by any camp staff; therefore, this finding is speculative.

A comparison of all four children suggests gender was not a factor predicting variability of IRI scores. That is, both Child 2, female, and Child 4, male, demonstrated similar patterns of slight decline on Frustration and Listening Comprehension levels on for Passage Comprehension. Child 4 was 1-year, 9-months older than Child 2, suggesting age also was not a predictor of differences in IRI scores. Because all four children demonstrated slight improvement on the majority of WRAT4 subtests, the WRAT4 did not appear to predict IRI scores for any of the children. In addition, although Child 3 evidenced significant improvement on the IRI, he remained consistent on the sentence comprehension portion of the IRI, suggesting the measure of reading comprehension on the WRAT4 does not predict reading comprehension as measured on the IRI.

Myles et al. (2002) found that when examining types of questions answered by participants, children with Asperger's Disorder demonstrated a significantly greater ability to answer literal and fact-based questions rather than questions requiring more abstract reasoning ability (Myles et al. 2002). Results from the current study indicated participants had greater difficulty answering inference type questions than other types of questions, though this result was found only on the pretest administration of the IRI. Given the abstract nature of inference questions, the results of the pretest IRI are commensurate with the results obtained in the study conducted by Myles et al. (2002), though the results of the posttest are not. The change in error rates for the Inference type question was statistically significant, and Cohen's (1988) guidelines suggested this improvement represented a medium effect. No other question type evidenced statistically significant changes from pre- to posttest.

The finding that participants improved in their abilities to answer inference questions over the course of the camp suggests that academic remediation for children with autism may result in greater ability use inferential skills. Increasing inference abilities was not a specified objective of the academic remediation program; however, classroom activities used this skill frequently. It is likely the Language Arts class, which focused on reading and comprehending stories, played a role in this increased skill. It is also possible individual tutoring may have increased this skill as reading comprehension was a focus of many of the individual tutoring plans. However, the behavioral plans of many of the children focused on social skills, and these plans also may have increased inference abilities. For example, "What do you think will happen to buildings when the wind from a tornado reaches 300 mph?" Such questions require the participant to use inferential thinking when responding, therefore indirectly practicing this skill. It is possible this practice was then generalized to reading comprehension. In addition, tutoring

often focused on reading comprehension skills, which include inference skills, and this may have further strengthened the inferential skills for participants.

Implications

The present study provides several implications for the academic remediation of children with autism using a summer camp format. For example, although the camp failed to produce significant gains in achievement among children with autism, a significant loss of academic achievement skills was not found, suggesting maintenance of academic skills. The camp's format of four classes with disguised curriculum and individual tutoring appeared effective in buffering summer learning loss and made slight, though not statistically significant, improvement in the achievement skills for nearly all participants.

Classroom instruction involved group activities lead by a teacher. Each student received assistance and support from counselors. The relationship of two classes – Mathematics and Language Arts – to the material measured on the WRAT4 and IRI is readily apparent. The use of disguised curriculum in these classes meant that materials were not taught in the classes at the camp in a traditional format. Therefore, the slight gains experienced by the participants suggested generalization from a nontraditional format to standardized testing. It also implies these classes were useful in maintaining skills in these areas. Science was not measured on the WRAT4 specifically, though the science class occasionally incorporated elements of Language Arts, such as reading books about Tornadoes, and Mathematics, such as counting out the number of materials needed to complete a project. The incorporation of science in the camp likely strengthened the generalization of Language Arts and Mathematics skills to improvement the WRAT4. Although spelling was not specifically addressed in a separate course in the summer camp, whenever students wrote for classes, counselors encouraged the use of correct spelling and

corrected misspelled words. Spelling also was included in individual tutoring plans, and repeated whole word exposure to words in Language Arts also may have lead to maintenance in this area.

The maintenance of skills for all children in camp implies academic remediation via summer camp programming will be effective in maintaining academic skills for children with a variety of autism spectrum diagnoses and functional levels. Additionally, the use of a disguised curriculum is important when discussing maintenance of skills, as the use of the disguised curriculum allowed for a traditional summer camp experience while apparently still allowing a buffer against summer learning loss. This implies summer learning loss can be prevented, and perhaps slight gains in achievements can be made, in children with autism without sacrificing the enjoyable and engaging experiences of a summer camp.

The findings of reliable scores obtained on the WRAT4 imply the WRAT4 can be a reliable measure of achievement skills in children with autism. Given the concise nature of the WRAT4 and the behavioral and interaction deficits that are commonly displayed by children with autism, this indicates the WRAT4 may be an appropriate substitute for screening academic abilities in children with autism in lieu of longer, more detailed achievement tests.

The findings of unreliable scores obtained on the IRI indicated the IRI is not a reliable measure of reading comprehension for children with autism. Therefore, teachers and practitioners may reconsider using an IRI to measure reading levels of children with autism, despite the face validity of the IRI. However, as noted previously, the behaviors typically seen among children with autism may have contributed to inconsistent results found on the IRI. In addition, although all examiners were adequately trained to administer the IRI, the administration of the IRI requires more subjectivity in determining correct answers than the WRAT4 (which

states clearly what answers are correct and what answers are not), and this may have contributed to the inconsistent IRI results.

Results from the comparison of the WRAT4 scores and IRI scores for four individual children suggested the WRAT4 is a poor predictor of reading level on the IRI and also was a poor predictor of change in performance on the IRI over time. Notably, the Sentence Comprehension test included on the IRI, which measured reading comprehension, also did not appear to be a good predictor of change on reading comprehension on the IRI. Because the IRI contains longer passages that have greater face validity with classroom textbooks and storybooks, this implies children with autism may demonstrate reading levels on longer passages that are inconsistent with their reading level as measured on the WRAT4. In addition, children with autism may have attention difficulties due to behaviors commonly seen in autism, such as self-stimulating behaviors and thinking about focal interests, and this may have negatively affected their ability to read and comprehend longer written passages.

Limitations

The current study presents several limitations that should be carefully considered. A limitation of the current study is that it did not include a control group for comparison. A control group would make it easier to determine the presence of summer learning loss or gains in achievement among matched controls that may have differed from the results of the participants in the study. However, the use of a control group would require a group of children with autism not receive potentially effective treatment, and this situation is considered unethical (APA, 2002). As the summer camp required the use of the majority of the summer, there was not enough time to allow for a waitlist group to be used as a control group. In addition, it would not be ethical or pragmatic to deny children in the control group access to other forms of therapy,

such as Speech and Language Therapy, which would have an unknown effect on the academic achievement of the children in the control group.

Beyond ethical concerns, given the pervasive nature of autism (APA, 2000), the use of typically developing peers as a control would be inappropriate for children with autism. In addition, given the relatively low prevalence rates of autism (APA, 2000; Fombonne, 2005), finding developmentally appropriate matches for each participant with autism would be extremely difficult. Similarly, the current study did not control for cognitive ability within the sample. Because cognitive abilities directly influence the ability of individuals to learn academic skills, it is unknown if the results of the present study can be applied to children with autism who have cognitive abilities within specific ranges (e.g., mental retardation, giftedness).

Another limitation related to the sample is the small sample size. Smaller samples reduce the power of tests such as MANOVA, which may have been better able to demonstrate small increases in achievement if a larger sample size were utilized. However, small sample sizes found in experimental research with children with autism are likely a reflection of the low prevalence rate of autism (APA, 2000; Fombonne, 2005). Small sample sizes are indeed typical for experimental studies focusing on children with autism (Frith & Snowling, 1983; Hung & Thelander, 1973; Lopata et al. 2008; Lopata et al. 2006; Minshew, et al. 1992; Nation et al. 2006). Studies that include 40 or more participants tend to be retrospective in nature and are typically chart reviews (Church et al. 2000; Mayes & Calhoun, 2003a; Mayes & Calhoun, 2003b).

An Additional limitation regarding the sample size was that the majority of participants had been diagnosed with autism rather than PDD-NOS or Asperger's Disorder. Thus, the results of the current study may not be applicable to children diagnosed with PDD-NOS or Asperger's

Disorder. Children with Asperger's Disorder demonstrate greater cognitive and language skills than children with Autism (APA, 2000); therefore, it is possible academic remediation, particularly remediation provided over the relatively (compared to a school year) short term of a summer camp may be more successful for children with Asperger's Disorder than children with Autism. The diagnosis of PPD-NOS merely requires a child to demonstrate many symptoms of autism, but may be diagnosed in children who do not meet all the requirements for autism. Therefore, this category may include a wide variety of behaviors and deficits compared to children diagnosed with Asperger's Disorder or Autism and therefore further study within the population of children diagnosed with PDD-NOS would be required to determine effectiveness of academic remediation.

The current study involved 19 (90%) participants who were identified as Caucasian by their parents, and the results of this study may not apply to children with autism who do not identify as Caucasian. For example, although the present study gathered information about parental educational achievement, but it did not gather information about the education or SES status of extended family members, who may play a greater role in parenting in non-Caucasian households compared to Caucasian households. Also, special educational services, the length of the school day, and the length of the school year may sometimes differ greatly in other cultures, and therefore the application of the results of this study may not generalize to children from other multicultural backgrounds or who are enrolled in educational systems in other countries. For example, students who come from school systems who have significantly shorter days may find a full day of academic remediation difficult, and a summer camp program in such a culture may need to utilize a shorter day. Further, some cultures may have different school calendars from that of the United States and the summer learning loss reported by Alexander et al. (2007) may

not occur in cultures that have shorter summer breaks. If summer breaks are shorter, another program besides a summer camp may be the most appropriate means of providing academic remediation.

The age limitations used at the camp may limit the generalization of results. That is, preschoolers and adolescents older than 13 years of age were not included in this study, and generalization of these results may not be appropriate for those populations. Similarly, children who were unable to complete standardized testing due to behavioral problems or lack of verbal ability were not included in the present study, and generalization of the results of this study to those populations also would be inappropriate. In addition, the participants were not a random sample of the United States as they all resided within in an hour drive of the camp.

The length of the camp also may have been a limitation of the study. While a significant period of time, eight weeks may not be long enough to affect significant change in academic skills. Many of the children demonstrated slight gains in achievement on the WRAT4; it remains unknown if these children would have made significant gains on the WRAT4 if the duration of the camp were longer and they had received more academic remediation.

The consistent results produced by participants on the WRAT4 indicated it was a valid assessment tool. However, the use of the IRI produced more variable results, and the use of the IRI may have been a limitation in the present study due to the lack of psychometric properties inherent in using an informal assessment measure. The behavioral problems inherent in autism also may have limited the abilities of the participants to complete the IRI. In addition, the inexperience of the examiners administering the IRI may have resulted in differences in scoring the IRI.

Suggestions for Future Research

The present study provides valuable information regarding assessment of achievement skills in children with autism and the effectiveness of summer camp programming to deliver academic remediation services for children with autism. However, the limitations of the current study suggest directions future research could take to further add to the literature in these areas.

Primarily, the limitations of the study are regarding the sample. Future research should replicate this study with a larger sample of children. As this study included primarily children diagnosed with autism, future research should replicate this study with children who have been diagnosed with PDD-NOS and Asperger's Disorder to determine if the assessment and intervention procedures used in this study are equally effective with those populations. It should be noted that given the heterogeneity of these disorders, future research likely would benefit from having more children with the same diagnoses and similar functional levels likely would allow for more accurate generalization of results.

Also, future research may seek to replicate this study with children who have more diversity regarding ethnicity and culture, to determine if such assessment and intervention methods are applicable to those populations. Lastly, the age restrictions placed on the sample indicate the results may not be applicable to children younger than 6 years of age or older than 13 years of age. Future research may seek to replicate this study with these populations, though modifications may need to be made; for example, preschool children may require a shorter day and a higher camper-to-counselor ratio.

The variability of results found using the IRI may have been the result of the lack of standardization and psychometric properties of the instrument rather than true variability in the performance of the participants, as well as behavioral problems exhibited by participants that are

common to autism and the inexperience of the examiners. Future research may seek to measure reading comprehension on longer passages by using another instrument, or test from another achievement assessment, that has better psychometric properties to more accurately measure this skill.

Similarly, the present study did not control for cognitive ability among participants, which may influence the effectiveness of the academic remediation. Future research may benefit by controlling for the cognitive ability of the participants. Children with lower cognitive abilities may require more repetition to gain mastery of concepts, and this may require longer class periods or a lengthier summer camp term than was employed by this study. Children with lower cognitive abilities may struggle more than their peers with higher cognitive abilities to generalize from a nontraditional classroom setting to a more traditional paper-and-pencil task. Changes over time for children with lower cognitive abilities may therefore be smaller than gains made by their peers with higher cognitive ability as measured by standardized assessments, though still carry significant meaning and benefit. Therefore, future research focusing on children with lower cognitive abilities may wish to employ a different method of achievement assessment. Children with higher cognitive abilities may experience learning gains in shorter time frames or with more abstract materials than their less cognitively developed peers. Children of higher cognitive abilities may experience frustration when participating in a curriculum that may not be challenging for them, and this may contribute to behavioral problems that may negatively impact the academic gains for all children at the camp.

The WRAT4 is typically considered a screener of academic ability and is comprised of a limited number of tests. Future research may benefit from more detailed assessment of the participants, such as administering a traditional, standardized achievement battery to all

participants. This additional assessment may reveal significant changes in a more specific skills area not measured by the WRAT4; for example, increases in phonetic skills or reading fluency, which were not measured by the WRAT4.

References

- Alexander, K. L., Entwisle, D. R. & Olson, L. S. (2007). Lasting consequences of the summer learning gap. *American Sociological Review*, 72(2), 167-180.
- Allington, R. L. (1978). Teacher ability in recording oral reading performance. *Academic Therapy*, 14, 187-192.
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders: DSM-IV-TR*. Washington, DC: Author.
- American Psychiatric Association. (2002). *Ethical Principles of Psychologists and Code of Conduct*. Washington, DC: Author.
- Anderson, W. W. (1977). Informal reading inventories: Commercial or conventional. *Reading World*, 17, 64-68.
- Applegate, M. D., Quinn, K. B., & Applegate, A. J. (2002). Levels of thinking required by comprehension questions in informal reading inventories. *The Reading Teacher*, 56(2), 174-180.
- Ashendorf, L., Jefferson, A. L., Green, R. C. & Stern, R. A. (2009). Test-retest stability on the WRAT-3 reading subtest in geriatric cognitive evaluations. *Journal of Clinical and Experimental Neuropsychology*, 31(5), 605-610.
- Baker, B. L. (1972). Camp Freedom: Behavior modification for retarded children in a therapeutic camp setting. *American Journal of Orthopsychiatry*, 43(3), 418-427.
- Bartak, L. & Rutter, M. (1973). Special educational treatment of autistic children: A comparative study—1. Design of study and characteristics of units. *Journal of Child Psychology & Psychiatry & Allied Discipline*, 14(3), 161-179.

Bateman, B. (1968). A pilot study of mentally retarded children attending summer day camp.

Mental Retardation, 6(1), 39-44

Behavior Characteristics Progression. Palo Alto, CA: VORT Corporation, 1973

Bennett, L. W., Cardone, S., Jarczyk, J. (1998). Effects of a therapeutic camping program on addiction recovery. *Journal of Substance Abuse Treatment*, 15(5), 469-474.

Berger, A. & Kautz, C. (1967). The Braille informal reading inventory. *The Reading Teacher*, 21(2), 149-152.

Betts, E. A. (1954). *Foundation of Reading Instruction*. New York: American Book Company.

Blachowicz, C. L. Z. (1999). Vocabulary in dynamic reading assessment: Two case studies. *Journal of Reading Psychology*, 20, 213-236.

Blake, J. (1996). Opening doors: Integration of persons with a disability in organized children's camping in Canada. *Journal of Leisurability*, 23(2), 3-10.

Blas, H. I. (2007). A specialized program for children with developmental disabilities within a "typical" overnight summer camp: Camp Ramah's Tikvah program. *Child and Adolescent Psychiatric Clinics of North America*, 16, 875-886.

Boomer, L. W. & Garrison-Harrell, L. (1995). Legal issues concerning children with autism and pervasive developmental disabilities. *Behavioral Disorders*, 21(1), 53-61.

Bradley, J. M. & Ames, W. S. (1976). The influence of intrabook readability variation on oral reading performance. *The Journal of Educational Research*, 70(2), 101-105.

Brecht, R. D. (1977). Testing format and instructional level with the informal reading inventory. *The Reading Teacher*, 31(1), 57-59.

Briery, B. G. & Rabian, B. (1999). Psychosocial changes associated with participation in a pediatric summer camp. *Journal of Pediatric Psychology*, 24(2), 183-190.

- Brookman, L., Boettcher, M., Klein, E., Openden, D., Koegel, R. L., & Koegel, L. K. (2003). Facilitating social interactions in a community summer camp setting for children with autism. *Journal of Positive Behavior Interventions*, 5(4), 249-252.
- Brooks, C. R. (1977). WISC, WISC-R, S-B L&M, WRAT: Relationships and trends among children ages six to ten referred for psychological evaluation. *Psychology in the Schools*, 14(1), 30-33.
- Bullock, C. C., Mahon, M. J. & Welch, L. K. (1992). Easter Seals' progressive mainstreaming model: Options and choices in camping and leisure services for children and adults with disabilities. *Therapeutic Recreation Journal*, 26(4), 61-70.
- Burns, P. C. & Roe, B. D. (2002). *Informal reading inventory* (6th ed.). Boston: Houghton Mifflin Company.
- Church, C., Alisanski, A., & Amanullah, S. (2000). The social, behavioral, and academic experiences of children with Asperger syndrome. *Focus on Autism and Other Developmental Disabilities*, 15(1), 12-20.
- Coonrod, E. E. & Stone, W. L. (2005). Screening for autism in young children. In Fr. Volkmar, R. Paul, A. Klin, and D. Cohen (Eds.), *Handbook of Autism and Pervasive Developmental Disorders: Vol. 2. Assessment, Interventions, and Policy* (3rd ed., pp. 42-69). Hoboken, NJ: John Wiley & Sons, Inc.
- Cooper, D. & Fraboni, M. (1988). Relationship between the Wechsler Adult Intelligence Scale-Revised and the Wide Range Achievement Test-Revised in a sample of normal adults. *Educational and Psychological Measurement*, 48, 799-803.
- Cooper, H. Charlton, K., Valentine, J. C. & Muhlenbruck, L. (2000). Making the most of summer school: A meta-analytic and narrative review. *Monographs of the Society for*

Research in Child Development, 65(1), 1-127

Cooper, H., Nye, B., Charlton, K., Lindsay, J. & Greathouse, S. (1996). The effects of summer vacation on achievement test scores: A narrative and meta-analytic review. *Review of Educational Research*, 66(3), 227-268.

Davis, E. E. (1975). Reading frustration level as indicated by polygraph. *The Journal of Educational Research*, 68(8), 286-288.

Davis-Berman, J. & Berman, D. S. (1989). The wilderness therapy program: An empirical study of its effects with adolescents in an outpatient setting. *Journal of Contemporary Psychology*, 19(4), 271-281.

Davis-Berman, J. & Berman, D. S. (1994). *Wilderness therapy: Foundations, theory & research*. Dubuque, IA: Kendall/Hunt Publishing Company.

Davis, E. E. & Ekwall, E. E. (1976). Mode of perception and frustration in reading. *Journal of Learning Disabilities*, 9(7), 53-59.

Dell, C. A., Harrold, B., Dell, B. (2008). Test Review. *Rehabilitation Counseling Bulletin*, 52(1), 57-60.

Edwards, T., Kahn, S. & Brenton, L. (2001). Math corps summer camp: An inner city intervention program. *Journal of Education for Students Placed at Risk*, 6(4), 411-426.

Eikenaes, I., Gude, T., & Hoffart, A. (2006). Integrated wilderness therapy for avoidant personality disorder. *Nordic Journal of Psychiatry*, 60(4), 275-281.

Ekwall, E. E. & Shanker, J. L. (1993). *Ekwall/Shanker Reading Inventory – Third Edition*. Boston: Allyn and Bacon.

Ekwall, E. E., Solis, J. K. E. & Solis, E. (1973). Investigating informal reading inventory scoring criteria. *Elementary English*, 50(2), 271-274, 323.

- Ekwall, E. E. (1974). Should repetitions be counted as errors? *The Reading Teacher*, 27(4), 365-367.
- Eng, B. & Davies, B. (1991). Effects of a summer camp experience on self-concept of children with cancer. *Journal of Pediatric Oncology Nursing*, 8(2), 89-90.
- Field, A. (2005). *Discovering Statistics Using SPSS (2nd ed.)*. London: Sage Publications.
- Fombonne, E. (2005). Epidemiological studies of Pervasive Developmental Disorders. In Fr. Volkmar, R. Paul, A. Klin, and D. Cohen (Eds.), *Handbook of Autism and Pervasive Developmental Disorders: Vol. 1. Diagnosis, Development, Neurobiology, and Behavior* (3rd ed., pp. 42-69). Hoboken, NJ: John Wiley & Sons, Inc.
- Forehand, R., Mulhern, T., Rickard, H. C. (1969). Effects of token reinforcement in a therapeutic camp. *Psychological Reports*, 25, 349-350.
- Frith, U. & Snowling, M. (1983). Reading for meaning and reading for sound in autistic and dyslexic children. *British Journal of Developmental Psychology*, 1, 329-342.
- Fuchs, L. S., Fuchs, D., & Deno, S. L. (1982). Reliability and validity of curriculum-based informal reading inventories. *Reading Research Quarterly*, 18(1), 6-26.
- Gagnon, L., Mottron, L. Bherer, L. & Joanne, Y. (2004). Quantification judgment in high functioning autism: Superior or Different? *Journal of Autism & Developmental Disorders*, 34(6), 679-689.
- Gerke, R. (1980). Critique of informal reading inventories: Can a valid instructional level be obtained? *Journal of Reading Behavior*, 12(2), 155-158.
- Gillis, H. L., & Simpson, C. (1991). Adventure-based residential drug treatment for court-referred youth. *Journal of Addictions & Offender Counseling*, 12(1), 12.
- Glass, J. S. & Myers, J. E. (2001). Combining the old and the new to help adolescents:

- Individual psychology and adventure based counseling. *Journal of Mental Health Counseling*, 23(2), 104-114.
- Goldstein, G., Minshew, N. J. & Siegel, D. J. (1994). Age differences in academic achievement in high-functioning autistic individuals. *Journal of Clinical and Experimental Neuropsychology*, 16(5), 671-680.
- Gonzales, P. C., & Elijah, D. V. (1975). Rereading: Effect on error patterns and performance levels on the IRI. *The Reading Teacher*, 28(7), 647-652.
- Goodwin, D. L., & Staples, K. (2005). The meaning of summer camp experiences to youth with disabilities. *Adapted Physical activity quarterly*, 22, 160-178.
- Grigorenko, E. L., Klin, A., Pauls, D. L., Senft, R., Hooper, C. & Volkmar, F. (2002). A descriptive study of hyperlexia in a clinically referred sample of children with developmental delays. *Journal of Autism & Developmental Disorders*, 32(1), 3-12.
- Grigorenko, E. L., Klin, A. & Volkmar, F. (2003). Annotation: Hyperlexia: Disability or superability? *Journal of Child Psychology & Psychiatry & Allied Disciplines*, 44(8), 1079-1091.
- Griswold, D. E., Barnhill, G. P., Myles, B. S., Hagiwara, T., & Simpson, R. L. (2002). Asperger Syndrome and academic achievement. *Focus on Autism and Other Developmental Disabilities*, 17(2), 94-102.
- Hammill, D. D. (1985). *DTLA – 2: Detroit Tests of Learning Aptitude*. Austin, TX: Pro-Ed.
- Heckel, R. V., Hursh, L., & Hiers, J. M. (1977). Analysis of process data from token groups in a summer camp. *Journal of Clinical Psychology*, 33(1), 241-244.
- Heflin, L. J. & Alaimo, D. F. (2007). *Students with Autism Spectrum Disorders*. Upper Saddle River, NJ: Pearson Education, Inc:

- Heflin, L. J., & Simpson, R. L. (1998). Interventions for children and youth with autism: Prudent choices in a world of exaggerated claims and empty promises. Part I: Legal/policy analysis and recommendations for selecting interventions and treatments. *Focus on Autism and Other Developmental Disabilities*, 13(4), 212-220.
- Helgren-Lempesis, V. A., & Mangurm, C. T. (1986). An analysis of alternate-form reliability of three commercially-prepared informal reading inventories. *Reading Research Quarterly*, 21(2), 209-215.
- Henley, D. (1999). Facilitating socialization within a therapeutic camp setting for children with attention deficits utilizing the expressive therapies. *American Journal of Art Therapy*, 38, 40-50.
- Herr, D. E. (1977). Institutionalized adolescents' perceptions of a summer camp program. *Adolescence*, 12(47), 421-431
- Heyns, B. (1978). *Summer learning and the effects of schooling*. New York: Academic Press.
- Hinshaw, S. P. (2002). Preadolescent girls with Attention-Deficit/Hyperactivity Disorder: I. Background characteristics, comorbidity, cognitive and social functioning, and parenting practices. *Journal of Consulting and Clinical Psychology*, 70(5), 1086-1098.
- Hung, D. W. (1977). Generalization of 'curiosity' questioning behavior in autistic children. *Journal of Behavior Therapy and Experimental Psychiatry*, 8(3), 237-245.
- Hung, D. W., & Thelander, M. J. (1978). Summer camp treatment for autistic children. *Exceptional Children*, 44(7), 534-536.
- Jastak, J. F. & Jastak, S. R. (1965). *The Wide Range Achievement Test*. Wilmington DE: Guidance Associates.

- Jelalian, E., Mehlenbeck, R., Lloyd-Richardson, E. E., Birmaher, V. & Wing, R. R. (2006). 'Adventure Therapy' combined with cognitive-behavioral treatment for overweight adolescents. *International Journal of Obesity*, 30(1), 31-39.
- Jerrolds, B. W., Callaway, B. & Gwaltney, W. (1971). A comparative study of three tests of intellectual potential, three tests of reading achievement, and the discrepancy scores between potential and achievement. *The Journal of Educational Research*, 65(4), 168-172.
- Johns, J. (1983). The informal reading inventory: 1910-1980. *Reading World*, 23(1), 8-19.
- Johnson, M. S., Kress, R. A., & Pikulski, J. J. (1987). Informal Reading Inventories. Newark: DE: International Reading Association.
- Jones, C. D. (2004). The effectiveness of wilderness adventure therapy programs for young people involved in the juvenile justice system. *Residential Treatment for Children & youth*, 22(2), 53-67.
- Karnes, F. A., Edwards, R. P., & McCallum, R. S. (1986). Normative achievement assessment of gifted children: Comparing the K-ABC, WRAT, and CAT. *Psychology in the Schools*, 23, 346-352.
- Kaufman, A. S. & Kaufman, N. L. (1985). *Kaufman Test of Educational Achievement*. Circle Pines, MN: American Guidance Service.
- Kaufman, A. S. & Kaufman, N. L. (2004). *Kaufman Test of Educational Achievement – Second Edition*. Circle Pines, MN: American Guidance Service.
- Kelly, M. P., Coursey, R. D., & Selby, P. M. (1997). Therapeutic adventures outdoors: A demonstration of benefits for people with mental illness. *Psychiatric Rehabilitation Journal*, 20(4), 61.

- Kim, J. (2004). Summer reading and the ethnic achievement gap. *Journal of Education For Students Placed at Risk*, 9(2), 169-188.
- Klee, K., Greenleaf, K. & Watkins, S. (1997). Summer camps for children and adolescents with kidney disease. *American Nephrology Nurses Association Journal*, 24(1), 57-61.
- Klesius, J. P. & Homan, S. P. (1985). A validity and reliability update on the informal reading inventory with suggestions for improvement. *Journal of Learning Disabilities*, 18(2), 71-76.
- Klimczak, N. C., Bradford, K. A., Burright, R. G., & Donovanick, P. J. (2000). K-FAST and WRAT-3: Are they really different? *The Clinical Neuropsychologist*, 14(1), 135-138.
- Larson, B. A. (2007). Adventure camp programs, self-concept, and their effects on behavioral problem adolescents. *Journal of Experiential Education*, 29(3), 313-330.
- Lattal, K. A. (1969). Contingency management of toothbrushing behavior in a summer camp for children. *Journal of Applied Behavior Analysis*, 2, 195-198.
- Leviton, H. & Kiraly, J. (1975). Summer compensatory education: Effects on learning disabled children. *Exceptional Children*, 42(1), 48-49.
- Lopata, C., Thomeer, M. L., Volker, M. A. & Nida, R. E. (2006). Effectiveness of a cognitive behavioral treatment on the social behaviors of children with Asperger's Disorder. *Focus on Autism & Other Developmental Disabilities*, 21(4), 237-244.
- Lopata, C., Thomeer, M. L., Volker, M. A., Nida, R. E., & Lee, G. K. (2008). Effectiveness of a manualized summer social treatment program for high-functioning children with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 38, 890-904.

- Lord, C. (1995). Follow-up of two year olds referred for possible autism. *Journal of Child Psychology and Psychiatry*, 36(8), 1365-1382.
- Lovitt, T. C. & Hansen, C. L. (1976). Round one – placing the child in the right reader. *Journal of Learning Disabilities*, 9(6), 18-24.
- Luckner, J. L. (1989). Altering locus of control in individuals with hearing impairments by outdoor-adventure courses. *Journal of Rehabilitation*, 55(2), 62-67.
- Mandlawitz, M. R. (2005). Educating Children with Autism: Current Legal Issues. In Fr. Volkmar, R. Paul, A. Klin, and D. Cohen (Eds.), *Handbook of Autism and Pervasive Developmental Disorders: Vol. 2. Assessment, Interventions, and Policy* (3rd ed., pp. 1161-1173). Hoboken, NJ: John Wiley & Sons, Inc.
- Marr, M. B. & Lyon, K. R. (1981). Passage independency and question characteristics: An analysis of three informal reading inventories. *Reading Psychology*, 2(2), 97-102.
- Martiniuk, A. L. C. (2003). Camping programs for children with cancer and their families. *Supportive Care in Cancer*, 11(12), 749-757.
- Marx, J. D. (1988). An outdoor adventure counseling program for adolescents. *Social Work*, 33(6), 517-520.
- Masztal, N. B. & Smith, L. L. (1984). Do teachers really administer IRIs? *Reading World*, 24(1), 80-83.
- Mayes, S. D. & Calhoun, S. L. (2003a). Ability profiles in children with autism. *Autism*, 6(4), 65-80.
- Mayes, S. D., & Calhoun, S. L. (2003b) Analysis of WISC-III, Stanford-Binet: IV, and academic achievement test scores in children with autism. *Journal of Autism and Developmental Disorders*, 33(3), 329-341.

- Mayes, S. D. & Calhoun, S. L. (2008). WISC-IV and WIAT-II profiles in children with high-functioning autism. *Journal of Autism and Developmental Disorders*, 38, 428-438.
- McCabe, P. P., Margolis, H., & Mackie, B. (1991). The consistency of reading disabled students' instructional levels as determined by The Metropolitan Achievement Test and the Ekwall Informal Reading Inventory. *Reading Research and Instruction*, 30(3), 53-62.
- McCabe, P. P., Margolis, H., & Barenbaum, E. (2001). A comparison of Woodcock-Johnson Psychoeducational Battery-Revised and Qualitative Reading Inventory-II instructional reading levels. *Reading & Writing Quarterly*, 17, 279-289.
- McIntosh, D. & Rawson, H. E. (1988). Effects of a structured behavior modification treatment program on locus of control in behaviorally disturbed children. *Journal of Genetic Psychology*, 149(1), 45-51.
- Merwin, J. C. (1972). Review of the Wide Range Achievement Test. In O. K. Buros (Ed.) *The seventh mental measurements yearbook*, (pp. 36-37). Highland Park, NJ: Gryphon.
- Michalski, J. H., Mishna, F., Worthington, C. & Cummings, R. (2003). A multi-method impact evaluation of a therapeutic summer camp program. *Child and Adolescent Social Work Journal*, 20(1), 53-76.
- Minshew, N. J., Goldstein, G., Muenz, L. R. & Payton, J. B. (1992). Neuropsychological functioning in nonmentally retarded autistic individuals. *Journal of Clinical and Experimental Neuropsychology*, 14(5), 749-761.
- Minshew, N. J., Goldstein, G., Taylor, H. G., & Siegel, D. J. (1994). Academic achievement in high functioning autistic individuals. *Journal of Clinical and Experimental Neuropsychology*, 16(2), 261-270.
- Myles, B. S., Hilgenfeld, T. D., Barnhill, G. P., Griswold, D. E., Hagiwara, T., & Simpson, R. L.

- (2002). Analysis of Reading Skills in Individuals with Asperger Syndrome. Focus on Autism and Other Developmental Disabilities, 17(1), 44-47.
- Myles, B. S., Huggins, A., Rome-Lake, M., Hagiwara, T., Barnhill, G. P., & Griswold, D. E. (2003). Written language profile of children and youth with Asperger syndrome: From research to practice. *Education and Training in Developmental Disabilities*, 38(4), 362-369.
- Naglieri, J. A. (1980). McCarthy and WISC-R Correlations with WRAT achievement scores. *Perceptual and Motor Skills*, 51, 392-394.
- Nation, K., Clarke, P., Wright, B., & Williams, C. (2006). Patterns of reading ability in children with autism spectrum disorder. *Journal of Autism & Developmental Disorder*, 36(7), 911-919.
- National Research Council. (2001). *Educating children with autism*. Committee on Educational Interventions for Children with Autism. Division of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.
- Newman, T. M., Macomber, D., Naples, A. J., Babitz, T., Volkmar, F. & Grigorenko, E. L. (2007). Hyperlexia in children with autism spectrum disorders. *Journal of Autism & Developmental Disorders*, 37(4), 760-774.
- Nilsson, N. L. (2008). A critical analysis of eight informal reading inventories. *The Reading Teacher*, 61(7), 526-536.
- Packman, W., Fine, J., Chesterman, B., vanZutphen, K., Golan, R., & Amylon, M. D. (2004). Camp Okizu: Preliminary investigation of a psychological intervention for siblings of pediatric cancer patients. *Children's Health Care*, 33(3), 201-215.

- Paris, S. G. (2002). Measuring children's reading development using leveled texts. *The Reading Teacher*, 56(2), 168-170.
- Paris, S. G. & Carpenter, R. D. (2003). FAQs about IRIs. *The Reading Teacher*, 56(6), 578-580.
- Peterson, J., Greenlaw, J. M., & Tierney, R. J. (1978). Assessing instructional placement with the IRI: The effectiveness of comprehension questions. *The Journal of Educational Research*, 71(5), 247-250.
- Pikulski, J. (1974). A critical review: Informal reading inventories. *The Reading Teacher*, 28(2), 141-151.
- Pikulski, J. (1990). Informal Reading Inventories. *The Reading Teacher*, 43(7), 514-516.
- Pikulski, J. J. & Shanahan, T. (1982). Informal reading inventories: A critical analysis. In J. J. Pikulski & T. Shanahan (Eds.), *Approaches to the informal evaluation of reading* (pp. 94-116). Newark, DE: International Reading Association.
- Piotrowski, C. & Keller, J. W. (1989). Psychological testing in outpatient mental health facilities: A national study. *Professional Psychology: Research and practice*, 20(6), 423-425.
- Powars, D. R. & Brown, M. (1990). Sick cell disease. Summer Camp. Experience of a 22-year community-supported program. *Clinical Pediatrics*, 29(2), 81-85.
- Prewett, P. N. Bardos, A., & Fowler, D. B. (1991). Relationship between the KTEA brief and comprehensive forms and the WRAT-R level 1 with referred elementary school students. *Educational and Psychological Measurement*, 51, 729-734.
- Prewet, P. N. & Giannuli, M. M. (1991). The relationship among the reading subtests of the WJ-R, PIAT-R, K-TEA, and WRAT-R. *Journal of Psychological Assessment*, 9, 166-174.
- Rasinski, T. V. (1999). Exploring a method for estimating independent, instructional, and frustration reading rates. *Journal of Reading Psychology*, 20, 61-69.

- Rawson, H. E. & McIntosh, D. E. (1991). The effects of therapeutic camping on the self-esteem of children with severe behavior problems. *Therapeutic Recreation Journal*, 25(4), 41-49.
- Reid, N. (1986). Wide Range Achievement Test: 1984 revised edition. *Journal of Counseling and development*, 64, 538-539.
- Reynolds, C. R., Dappen, L., Wright, D. (1981). A comparison of the criterion-related validity (academic achievement) of the WPPSI and the WISC-R. *Psychology in the Schools*, 18(20-23).
- Rice, C. (2007). Prevalence rates of autism spectrum disorders – autism and developmental disabilities monitoring network, 14 sites, United States 2002. *Morbidity and Mortality Weekly Report*, 56(1), 12-28.
- Rickard, H. C., Clements, C. B. & Willis, J. W. (1970). Effects of contingent and noncontingent token reinforcement upon classroom performance. *Psychological Reports*, 27, 903-908.
- Rickard, H. C., Taylor, G. E. & Libb, J. W. (1973). Contingency management and rate of classroom productivity. *Perceptual and Motor Skills*, 36, 1267-1273.
- Romi, S. & Kohan, E. (2004). Wilderness programs: Principles, possibilities and opportunities for intervention with dropout adolescents. *Child & Youth Care Forum*, 33(2), 115-136.
- Rumsey, J. M. & Hamburger, S. D. (1990). Neuropsychological divergence of high-level autism and severe dyslexia. *Journal of Autism and Developmental Disorders*, 20(2), 155-168.
- Russell, K. C. (2001). What is wilderness therapy? *The Journal of Experiential Education*, 24(2), 70-79.
- Russell, K. C. (2003). An assessment of outcomes in outdoor behavioral healthcare treatment. *Child & Youth Care Forum*, 32(6), 355-381.

- Rutter, M. & Bartak, L. (1973). Special educational treatment of autistic children: A comparative study-II. Follow-up findings and implications for services. *Journal of Child Psychology & Psychiatry & Allied Disciplines*, 14(4), 241-270.
- Rynders, J. E., Schleien, S. J. & Mustonen, T. (1990). Integrating children with severe disabilities for intensified outdoor education: Focus on feasibility. *Mental Retardation*, 28(1), 7-14.
- Sattler, J. M., & Ryan, J. J. (1981). Relationship between WISC-R and WRAT in children referred for learning difficulties. *Psychology in the Schools*, 18, 290-292.
- Schacter, J. (2003). Preventing summer reading declines in children who are disadvantaged. *Journal of Early Intervention*, 26(1), 47-58.
- Schacter, J. & Jo, B. (2005). Learning when school is not in session: A reading summer day-camp intervention to improve the achievement of exiting first grade students who are economically disadvantaged. *Journal of Research in Reading*, 28(2), 158-169.
- Schwarting, F. G., & Schwarting, K. R. (1977). The relationship of the WISC-R and WRAT: A study based upon a selected population. *Psychology in the Schools*, 14(4), 431-433
- Shasby, G., Heuchert, C. & Gansneder, B. (1984). The effects of a structured camp experience on locus of control and self concept of special populations. *Therapeutic Recreation Journal*, 18(2), 32-40.
- Smith, T. D., Smith, B. L. (1983). Relationship between the Wide Range Achievement Test 3 and the Wechsler Individual Achievement Test. *Psychological Reports*, 83, 963-967.
- Smith, T. C., Smith, B. L. & Dobbs, K. (1991). Relationship between Peabody Picture Vocabulary Test-Revised, Wide Range Achievement Test-Revised, and Weschler Intelligence Scale for Children-Revised. *Journal of School Psychology*, 29, 53-56.

- Smith, T. D., Smith, B. L., Smithson, M. M. (1995). The relationship between the WISC-III and the WRAT 3 in a sample of rural referred children. *Psychology in the Schools*, 32, 291-295.
- Snelbacker, A. J., Wilkinson, G. S., Robertson, G. J. & Glutting, J. J. (2001). Wide Range Achievement Test 3 (WRAT3). In W. I. Dorfman and M. Hersen (Eds.), *Understanding Psychological Assessment* (pp.259-317). New York: Kluwer Academic/Plenum Publishers.
- Spector, J. R. (2005). How reliable are informal reading inventories? *Psychology in the Schools*, 42(6), 593-603.
- Spruill, J. & Beck, B. (1986a). Relationship between the WIAS-R and Wide Range Achievement Test-Revised. *Educational and Psychological Measurement*, 46, 1037-1040.
- Spruill, J. & Beck, B. (1986b). Relationship between the WRAT and WRAT-R. *Psychology in the Schools*, 23, 357-360.
- Stevens, B. (2004). Adventure therapy for adolescents with cancer. *Pediatric Blood & Cancer*, 43(3), 278-284.
- Stevenson, Jr., J. D., (1990). A comparison of the 1978 and 1984 forms of the Wide Range Achievement Test with the WAIS-R for rehabilitation agency adults. *Educational and Psychological Measurement*, 50, 403-410.
- Stevens, J. P. (1980). Power of the multivariate analysis of variance tests. *Psychological Bulletin*, 88, 728-737.
- Stich, T. F. & Senior, N. (1984). Adventure therapy: An innovative treatment for psychiatric patients. *New directions for mental health services*, 21(1), 103-108.

- Sullivan, T. E., Hawkins, K. A. (1995). Support for abbreviation of the wide range achievement test-revised spelling subtest in neuropsychological assessments. *Journal of Clinical Psychology, 51*(4), 552-554
- Szatmari, P., Tuff, L., Finlayson, M. & Bartolucci, G. (1990). Asperger's syndrome and autism: Neurocognitive aspects. *Journal of the American Academy of Child and Adolescent Psychiatry, 29*(1), 130-136.
- Tabachnick, B. G. & Fidell, L. S. (2007). *Using multivariate statistics (5th ed.)*. Boston: Pearson Education, Inc.
- Valmont, W. J. (1972). Creating questions for informal reading inventories. *Reading Teacher, 25*(6), 509-512.
- Vance, B. & Fuller, G. B. (1995). Relation of scores on WISC-III and WRAT-3 for a sample of referred children and youth. *Psychological Reports, 76*, 371-374.
- Van Wert, M. & Reitz, A. L. (1978). A summer program for autistic children. *Education and Treatment of Children, 1*(2), 47-55.
- Vance, H. R., Hankins, N., & Brown, W. Predictive validity of the McCarthy Screening test based on Wide Range Achievement Test. *Psychological Reports, 59*, 1060-1062.
- Vance, B., Kitson, D., & Singer, (1985). Relationship between the standard scores of Peabody Picture Vocabulary Test-Revised and Wide Range Achievement Test. *Journal of Clinical Psychology, 41*(5), 691-693.
- Wazana, A., Bresnahan, M., & Kline, J. (2007). The autism epidemic: Fact or artifact? *Journal of the American Academy of Child & Adolescent Psychiatry, 46*(6), 721-730.

- Webster, R. E., Hewett, J. B., & Crumbacker, H. M. (1989). Criterion-related validity of the WRAT-R and K-TEA with teacher estimates of actual classroom academic performance. *Psychology in the Schools*, 26, 243-248.
- Wechsler, D. (2001). *Wechsler Individual Achievement Test, Second Edition*. San Antonio, TX: The psychological Corporation.
- Westervelt, V. D., Johnson, D. C., Westervelt, M. D. & Murrill, S. (1998). Changes in self-concept and academic skills during a multimodal summer camp program. *Annals of Dyslexia*, 48, 191-212.
- Whitehouse, D. & Harris, J. C. (1984). Hyperlexia in infantile autism. *Journal of Autism and Developmental Disorders*, 14(3), 281-289.
- Wilkinson, G. S., & Robertson, G. J. (2006). *Wide Range Achievement Test (4th ed.)*. Lutz, FL: Psychological Assessment Resources
- Witt, J. C., (1986). Review of the Wide Range Achievement Test-Revised. *Journal of Psychological Assessment*, 4, 87-90.
- Woodcock, R. W. (1987). *Woodcock Reading Mastery Tests-Revised*. Circle Pines, MN: American Guidance Service.
- Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock Johnson III Tests of Achievement*. Itasca, IL: Riverside Publishing.
- Yssel, N., Margison, J., Cross, T., & Merbler, J. (2005). Puzzles, mysteries, and Picasso: A summer camp for students who are gifted and learning disabled. *TEACHING Exceptional Children*, 38(1), 42-46.
- Zemke, R., Knuth, S. & Chase, J. (1984). Change in self-concepts of children with learning

difficulties during a residential camp experience. *Occupational Therapy in Mental Health*, 4(4), 1-12.

Zwaig, M. S. (1974). A successful camp experience for the LD child. *Academic Therapy*, 9(6), 445-449.

Appendix A

Appendix A

(Research Participant Application/Consent Form)

Research Participant Application Form

Exploring the Effectiveness of a Academic Remediation Summer Camp Program for Children with Autism Spectrum Disorders

Project Description: The purpose of this study is to investigate the effectiveness of an academic remediation program for children with autism spectrum disorders within a summer camp setting. This study will be a preliminary investigation into the application of basic academic instructional strategies designed to help children increase academic skills, which will facilitate the acquisition of new skills, within a summer camp setting. This project will be done over the summer of 2008.

Teresa Laking, doctoral student in school psychology, and the tutors/counselors for the students, will assess the reading, math, and written language skills for each child during the first day of camp and again during the last week of camp.

There is no current research on the effectiveness of academic instruction for children with autism in a summer camp setting. The National Research Council (2001) identified several characteristics of effective interventions for young children with autism, defined as 25 hours per week, 5 days per week, and 12 months per year. While many children receive such services during the school year, there is a lack of services provided during the summer months, especially for children who attend public schools. Given the need for intensive instructional programming throughout the year, it seems not only appropriate, but necessary to provide children with autism and their families with services during the summer months.

Confidentiality of Research Project: For research purposes, all information regarding you and your child will be kept confidential. Your and your child's name or other identifying information will not be released for any reason.

Requirements of participation: Criteria for inclusion in the study will be based on seven principles: (1) the child is diagnosed with autism or an autism spectrum disorder by an outside professional using criteria from the *Diagnostic and Statistical Manual of Mental Disorders – Fourth Edition – Text Revision (DSM-IV-TR;* American Psychiatric Association, 2000); (2) each child is 6 to 12 years of age; (3) each child is presenting with behavioral problems that adversely affect his current functioning in social, academic, or daily living; (4) no child exhibits severe self-injurious behavior; (5) no child is severely limited by a physical handicap so that the camping experience might be hindered; (6) no child is psychotic or brain damaged; (7) the child's parents, guardians, or referral agency is interested in participation in the program and research study and agrees to cooperate in providing full application information (including permission to assess the child for evaluation and/or research purposes), providing records information (i.e., child's school individualized education plan and most recent psychoeducational

report), answering a parent questionnaire, filling out two well known standardized behavior rating scales for their child (Gilliam Autism Rating Scale, Second Edition [GARS-2] and the Adaptive Behavior Assessment System, Second Edition [ABAS-II]), and providing transportation to and from the camp daily.

Potential Risks

There is minimal to no risk for the present study. Academic tasks included in both assessments are similar to tasks completed in a typical public school classroom, such as completing mathematics problems and reading sentences. Therefore, no more risk should be found in this study than found in a typical public school classroom. Both assessments allow for basal and ceiling rules which accommodate students of different achievement levels and prevent frustration. If at anytime a participant expresses discomfort with the assessments, the examiner will discontinue testing immediately. All materials will be destroyed after they are no longer needed for analysis.

Potential Benefits

Professionals will gain knowledge and skills in the application of academic remediation in the setting of a summer camp. In particular, knowledge will be gained about the continuation of academic instruction through the summer months.

Participation is Voluntary: You and your child's participation in the research project is voluntary. You and your child can withdraw from the project at any time without negative consequences and, if you request, all data pertaining to you and your child will be destroyed. Please feel free to ask questions of the investigators before signing the Informed Consent Form and agreeing to your child's participation in this study and any time during the study.

Questions? Please call Dr. McIntosh or Teresa Laking, Principle investigators, at (765)285-8500, with any questions. If you have any questions regarding the rights of research participants, please contact Melanie Morris Coordinator of Research Compliance, Office of Academic Research and Sponsored Programs, Ball State University, Muncie, IN 47306, (765)285-5070.

There are no incentives provided to parents who participate in this study. There will be no financial expenses to the parents that participate in the study.

Parent Signature

Date

Parent Signature

Date

David E. McIntosh, PhD., HSPP
Dept. of Special Education
Meeks Distinguished Professor
(765)285-8500
demcintosh@bsu.edu

Research Participant Information

Parent Name(s): _____

Address(s): _____

Home Phone: _____ Work Phone: _____

Email: _____

What is the best way for us to contact you? _____

When is the best time for us to contact you? _____

Highest grade level completed by each parent in the home (place and X where appropriate)

Father: _____	Less than 9 th grade	Mother _____	Less than 9 th grade
_____	Some high school, no diploma	_____	Some high school, no diploma
_____	High School diploma or GED	_____	High School diploma or GED
_____	1-3 years of college	_____	1-3 years of college
_____	4 or more years of college	_____	4 or more years of college

Current occupation:

Father: _____

Mother: _____

Child's Name: _____

Gender: M F

Date of Birth: _____

Age: _____

Ethnicity (circle):	African American	White (Caucasian)
	Hispanic/Latino	Asian
	Other: _____	

Does your child have non-communicative speech? _____

Please rate the severity of your child's autism:	Mild:	_____
	Moderate:	_____
	Severe:	_____

Does your child display any self-injurious behaviors? ____ Yes ____ No

(Children who display self-injurious behaviors will not be included in this study)

Does your child have any physical limitations that would hinder the camp experience?

____ Yes ____ No (Children who have physical limitations will not be included in this study)

Has your child been diagnosed with a psychotic disorder or brain damage? ____ Yes
____ No (Children with psychotic disorders or traumatic brain injury will not be included
in this study)

I have read and understand the above information and give permission for
_____ (insert child's name) to participate in the research
project entitled, "Exploring the Effectiveness of a Behavioral Remediation Summer Camp
Program for Children with Autism Spectrum Disorders"

Parent Signature

Date

Parent Signature

Date

David E. McIntosh, PhD., HSPP
Dept. of Special Education
Meeks Distinguished Professor
(765)285-8500
demcintosh@bsu.edu

Appendix B

IRBNet: Review Details

3/23/09 9:12 AM



Welcome to IRBNet
Teresa Laking

Study Manager
Create New Project

Project Administration
Project Overview
Designer
Share this Project
Sign this Package
Submit this Package
Delete this Package
Send Mail
Project History

Review Details

[89676-3] Exploring the Effectives of an Academic Remediation Day Camp on Increasing the Academic Skills of Ball State University IRB, Muncie, IN

Submission Details

Submitted To Ball State University IRB, Muncie, IN
Submitted by Teresa Laking
Submission Date 06/06/2008
Submission Type Revision
Local Board Reference Number

Review Details:

Agenda	Review Type	Status	Effective Date	Expiration Date
06/06/2008 01:45 PM	Expedited Review	Approved	06/06/2008	06/05/2009

Board Documents:

Document Type	Description	Last Modified	View
Approval Letter	Approval Letter	06/06/2008 03:55 PM	

Copyright © 2002-2009 Research Dataware. All Rights Reserved.